RISK PERCEPTION, JUDGMENT AND INFORMATION PROCESSING: EVIDENCE FROM EXPERIMENTS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Ying Cao

August 2012

RISK PERCEPTION, JUDGMENT AND INFORMATION PROCESSING: EVIDENCE FROM EXPERIMENTS

Ying Cao, Ph. D.

Cornell University 2012

This thesis investigates the role of cognition in the broader field of agricultural and food economics. Three studies in the thesis each targets on one aspect of human behavior.

Study 1 deals with non-standard belief. Using overconfidence and self-attribution, study 1 investigates how consumers react to temporary quality failure and perceive food safety risk depending on previous consumption experience. Results showed that people decrease the consumption when facing ambiguous signals regarding the food quality, but would not cease to eat altogether. Due to a taste of consistency, participants updated their risk perceptions and judgments based on their eating behaviors. Though experienced consumers picked up signals faster, their judgment was not better than those non-users due to a much stronger psychological bias.

Study 2 deals with non-standard decision-making. Using psychological terms such as cognitive dissonance and confirmation bias, study 2 reveals how individual consumers inadequately process information, pay limited attention to signals, and make decisions bias towards their initial choices. Results suggested that consumer's judgment and information processing depend a lot on their initial beliefs or consumption status. They used higher bidding prices to justify previous behaviors and selectively paid attention to information in favor of their initial choices. In terms of market responses, due to psychological biases among consumers, demand curves were less possible to shift down under food safety risk.

Study 3 deals with non-standard preference and its market responses. Results showed that asymmetric adverse incentive structure made a risk-averse loan officer inclined to reject loans to avoid risk of penalty. It provides *micro* explanations for *macro* level credit rationing phenomenon in the financial market. Lenders' reference-dependent utility under asymmetric adverse incentive structure broadens literature on institutional mechanism design to a behavioral scope.

BIOGRAPHICAL SKETCH

Ying (Jessica) Cao joined in the Charles H. Dyson School of Applied Economics and Management at Cornell as a PhD student in 2007. Prior to joining in Cornell, Jessica obtained with distinction a bachelor's and a master's degree in economics in Nankai University in 2005 and in Tsinghua University in 2007 respectively in China.

Jessica works extensively on behavioral analysis of individual decision-making involving risk and uncertainty. Her research primarily addresses issues of belief, judgment, information processing and the interaction with decision behaviors. The concentration of her research falls in the areas of *Development Economics*, *Agricultural Business and Finance*, and *Food and Marketing Analysis*. Recent research focuses on behavioral institutional design in banking and finance; non-standard belief, measures and impacts of overconfidence; and agricultural insurance choices and production decision under risk, etc.

Jessica recently received first prize of the 2012 *Institutional and Behavioral Economic Society* (IBES) paper competition under the *Agricultural and Applied Economic Association* (AAEA). She has also been awarded the *Richard D. Aplin Teaching Excellence Fund* three years in a row since 2009. Jessica serves as an anonymous referee for journals such as *Agricultural and Resource Economics Review* (ARER), *Agricultural Finance Review* (AFR), *Journal of Agricultural Business in Developing & Emerging Economies* (JADEE), and *China Economic Review* (CES), etc. Her works were accepted for publication at *China Agricultural Economic Review* and *Public Health and Nutrition*.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

First and foremost, I would like to express my deepest gratitude to my chair of special committee, Dr. David Just and Dr. Calum Turvey, for their excellent guidance, caring, patience, and providing me with an excellent atmosphere for doing research.

David offered me the opportunity for my very first piece of research on overconfidence at Cornell. He brought my attention to the realm of *Behavioral Economics*, and decision-making under risk in particular. He has been my inspiration to form innovative research ideas at the early stage.

Calum provided me with his unreserved trust and encouragement as I hurdled all the obstacles in the completion of this research work. He actively promoted me in the academic society; let me experience cutting-edge studies and practical issues in the field; patiently guided me through research design; and financially supported my research.

I would also like to thank my two other special committee members, Dr. William Schulze and Dr. Brian Wansink. Brian helped me to develop my background of consumer research in *Psychology* and *Marketing*. Bill gave me his best suggestions to improve my experimental design and the computer programs.

My gratitude is extended to the faculty members, staff, and graduate students in the Dyson School of Applied Economics and Management and many other related fields at Cornell. I would have been lonely without having them as good friends. Many thanks to Jiujie Ma, Rong Kong, Guangwen He, and workers in the Rural Credit Cooperatives in Shandong, China for helping me to coordinate the field study. My research would not have been possible without their helps.

Last but not least, I would like to thank my husband, Jiwei (Jerry) Zhao, and my parents. They were always supporting me and encouraging me with their best wishes. Especially, Jiwei was always there cheering me up and stood by me through the good times and bad.

TABLE OF CONTENTS

Contents

BIOGRAPHICAL SKETCH **iii**

ACKNOWLEDGEMENTS **iv**

LIST OF FIGURES **viii**

LIST OF TABLES **ix**

1 Chapter 1

Overview **1**

REFERENCES **5**

2 Chapter 2

Study 1: Consumer Belief and Risk Perception under Temporary Quality Failure **6**

2.1 Introduction 6

2.2 Literature Review 7

2.2.1 Cognitive Dissonance 7

2.2.2 Confirmation Bias 10

2.2.3 Consumer Behavior 12

2.3 Experimental Design 13

2.4 Experimental Results 15

2.4.1 Summary Statistics and Treatment Groups 15

2.4.2 Behavior 17

2.4.3 Perception 18

2.4.4 Dissonance 20

2.4.5 Future Behavior 21

2.5 Discussion 23

REFERENCES **25**

3 Chapter 3

Study 2: Cognitive Dissonance, Confirmation Bias and Inadequate Information Processing **31**

3.1 Introduction . . . 31

3.2 Experimental Design . . . 33

3.2.1 Treatments and Experimental Procedure . . . 33

3.2.2 Food Items, Information and Dissonance Inducing . . . 35

3.2.3 Auction Mechanism . . . 37

3.2.4 Testing Hypotheses . . . 40

3.3 Results . . . 44

3.3.1 Summary Statistics . . . 44

3.3.2 Cognitive Dissonance . . . 47

3.3.3 Manipulation Check . . . 49

3.3.4 Confirmation Bias . . . 51

3.3.5 Sticky Behavior . . . 52

3.3.6 Habits, Risk Perceptions and Causal Effects (Identification and Validation) . 54

3.3.7 Market Demand and Price Elasticity . . . 57

3.4 Discussion . . . 59

REFERENCES **62**

4 Chapter 4

Study 3: Credit Rationing under Lenders' Risk-Aversion and Asymmetric Adverse Incentives **65**

4.1 Introduction . . . 65

4.2 Literature Review . . . 69

4.3 Incentive Mechanisms and the Personal Responsibility System . . . 71

4.4 Theoretical Model . . . 73

4.4.1 Incentive Structures . . . 74

4.4.2 Reference-Dependent Utility . . . 75

4.4.3 Decision Strategies and Personal Equilibria . . . 77

4.4.4 Properties and Testing Hypotheses . . . 79
4.5 Experimental Design . . . 81
4.5.1 Loan Officers . . . 81
4.5.2 Loan Files . . . 81
4.5.3 Treatments . . . 82
4.5.4 Experimental Procedures . . . 84
4.5.5 Identification Strategy . . . 86
4.6 Results . . . 87
4.6.1 Summary Statistics . . . 87
4.6.2 Credit Rationing, Loan Decision and Accuracy . . . 89
4.6.3 Time Use, Confidence and Risk Aversion . . . 92
4.6.4 Credit Rating and Tendency to Approve . . . 95
4.7 Robustness . . . 98
4.7.1 Experimental Validity: Demographics . . . 100
4.7.2 Experimental Validity: Loan File Fixed Effects . . . 100
4.7.3 Experimental Validity: Order Effects . . . 101
4.7.4 Consistency . . . 102
4.8 Discussion . . . 102

REFERENCES **104**

5 Chapter 5

Conclusion **109**

APPENDIX **111**

List of Figures

1 Study 2_Risk Perception and Bid by Information Round & Treatment 47
2 Study 2_Demand Curves in Percentage Share - Peanut 59
3 Study 2_Demand Curves in Percentage Share - Almond 60
4 Study 2_Demand Curves in Percentage Share - Plain 60
5 Study 3_Kernel Density of Time Used per Loan 93
6 Study 3_Total Time Used per Loan across Rounds 94
7 Study 3_Percentage of Time Used by Part 94
8 Study 3_Tendency to Approve - Performing Loans 97
9 Study 3_Tendency to Approve - Non-Performing Loans 98
10 Study 3_Decision Accuracy Counts by Session 102

List of Tables

1 Study 1_Summay Statistics by Treatment 16

2 Study 1_Food Choice Behavior by Group 18

3 Study 1_Instrument for Behavior 19

4 Study 1_Impact of Food Choice on Perception 20

5 Study 1_Impact of Food Choice on Dissonance 21

6 Study 1_Impact of Food Choice on Future Behavior 22

7 Study 1_Participants Consuming None, Some and All of the Chicken by Group . . . 24

8 Study 2_Experimental Procedure 33

9 Study 2_Summary Statistics of Demographics 45

10 Study 2_Summary Statistics of Risk Perception and WTP by Round 45

11 Study 2_Average Treatment Effects on Bid - All & Preferred Only 48

12 Study 2_ATE on Bid - Robustness Check 49

13 Study 2_Independence of Risk Perception on Treatment & Preference 50

14 Study 2_ATE on Changes in Risk Perception - All & Preferred Only 50

15 Study 2_ATE on Changes in Bid - All & Preferred Only 53

16 Study 2_Preference Changes Before & After Study by Flavor 54

17 Study 2_ATE on Bid - Causal Inference 55

18 Study 2_ATE on Changes in Bid - Causal Inference, All 56

19 Study 2_ATE on Changes in Bid - Causal Inference, Preferred Only 57

20 Study 2_Inverse Demand Curve by Group 58

21 Study 3_Summary Statistics of Demographics 88

22 Study 3_Loan File Evaluation Frequency 89

23 Study 3_Final Payoffs by Treatment 89

24 Study 3_Frequency of Decision Accuracy 90

25 Study 3_ATE on Decision and Accuracy 91

26 Study 3_Marginal Effects of Time Use on Decision 96

27 Study 3_ATE on Tendency to Approve a Loan 99

1 Chapter 1

Overview

The classic economic theory builds on a simple but powerful model of behavior. It assumes individual agents are Bayesian information processors; have well-defined and stable preferences; maximize their expected utility; exponentially discount future well-being; are self-interested, narrowly defined; have preferences over final outcomes, not changes; have only "instrumental"/functional taste for beliefs and information, and are independent of the framing of the decision (Rabin, 2002).

Economic theories should be judged by three criteria: congruence with reality, generality, and tractability (Stigler, 1965). Over the past few decades, The research in Psychology and Economics (a.k.a. Behavioral Economics) suggests that individuals deviate from the standard model in three respects: (i) non-standard preferences; (ii) non-standard beliefs; and (iii) non-standard decision-making (DellaVigna, 2009). In the class of non-standard preferences, there are time preferences (self-control problems), risk preferences (reference dependence), and social preferences. On non-standard beliefs, there is evidence, for instance, on overconfidence, on the law of small numbers, and on projection bias. Regarding non-standard decision-making, there are i.e. framing, limited attention, menu effects, persuasion and social pressure, and emotions.

The historical development of Behavioral Economics presents 3 waves: First, identify anomalies—i.e., demonstrate clear violations of the assumption or model, and painstakingly rule out alternative explanations (such as subjects' confusion or transactions costs). Second, use the anomalies as inspiration to create alternative theories that generalize existing models. And third, construct economic models of behavior using the behavioral assumptions from the second step, derive fresh implications, and (empirically) test them.

As documented by Camerer and Loewenstein (2004), beginning around 1960, cognitive psychology became dominated by the metaphor of the brain as an information-processing device replacing the behaviorist conception of the brain as a stimulus-response machine. The information-processing metaphor permitted a fresh study of neglected topics like memory, problem solving and decision making. These new topics were more obviously relevant to the neoclassical conception of utility maximization than behaviorism had appeared to be. Psychologists such as Ward Edwards, Duncan Luce, Amos Tversky and Daniel Kahneman, began to use economic models as a benchmark against

which to contrast their psychological models.

The two most influential contributions were published by Tversky and Kahneman. Their 1974 article in Science argued that heuristic short-cuts created probability judgments which deviated from statistical principles. Their 1979 paper "Prospect theory: decision making under risk" documented violations of expected utility and proposed an axiomatic theory, grounded in psychophysical principles, to explain the violations. Later, during late 1980s and 1990s, more social scientists devoted to behavioral economics and made its development transited gradually from the first wave to the second wave. Rabin (1998, 2002) provide good survey of works within this period.

The recent past decade is a mixture of the second and the third wave. DillaVigna (2009) provides a comprehensive survey of the fruitful empirical research during this time. Different from previous work, DillaVigna's review was organized by psychological principles, rather than specific application topics. He also presented evidence on how rational actors - firms, employers, CEOs, investors, and politicians - respond to the non-standard behavior described in the survey; and finally, briefly discussed under what conditions experience and market interactions limit the impact of the non-standard features.

Studies of risk, perceptions and psychology at the cognitive scale have been very limited in the broader field of agricultural and food economics, yet in this area, and perhaps food in particular, the role of cognition could explain consumption behavior and demand characteristics beyond autarky. This thesis investigates the role of cognition across three subject areas, and establishes the reasonableness of the neoclassical model under imperfect or asymmetric market conditions.

In detail, in study 1, an experiment was conducted to investigate the interaction between consumers' past eating behaviors, risk perceptions and future information processing procedure. In the study, participants were required to choose whether or not to eat chicken that was potentially be tainted with Avian Influenza (AI). Results showed that people decreased the consumption when facing ambiguous signals regarding the food quality, but would not cease to eat altogether. Due to a taste of consistency, participants updated their risk perceptions and judgments based on their eating behaviors. The more chicken individuals ate the more favorably they tended to rate the food, suggesting confirmatory bias. Even though consumers with previous experience could pick up signals faster, their judgment was not better than those non-users due to a much stronger psychological bias. This study offered an explanation for why consumers were universally irresponsive to

public food safety information.

Study 2 investigated how individual consumers react to food safety information and make purchase decision. While it is expected that reading extra information about potential risk associated with food decreases consumers' willingness to pay (WTP), the magnitude of the impact varies across individuals. Based on findings in psychology, consumer's judgment and information processing may depend a lot on their initial beliefs or consumption status. Using an incentive compatible auction mechanism, this study elicited consumers' WTP under different informational settings. Results showed that consumers bid much higher when they freely chose food items (treatment) than when they were randomly assigned (control), suggesting cognitive dissonance. On average, the bidding premium was about 13 cents (roughly 30%) higher for low-risk food item and 30 cents (almost 60%) higher for high-risk item. The bidding premiums were further enlarged as food safety information was revealed to consumers. Confirmatory bias hypothesis was supported by the finding that free-choice group was more reluctant to change the bids despite of increased risk perceptions. In terms of market responses, due to psychological biases among consumers, demand curves were less possible to shift down under food safety risk. Results in this study implied that consumers were less responsive to public information due to their existing habits. Extra strategies would be needed to increase the efficiency of public communication to promote health.

In study 3, a field experiment was conducted to test loan decision behaviors. In this study the principal instrument for affecting risk perceptions are incentive mechanisms imposed by a regulatory authority. To what extent the incentives distort attitudes towards risk is also a problem in psychology, albeit one that has not been explored in detail previously. This study investigates how does incentive mechanism alter the perceptions of risk or the judgements about risk taking activities, i.e. Type I vs Type II error in lending decisions.

Study 3 showed that lenders' risk-aversion and behavioral responses resulted in credit rationing under certain incentive schemes. In the experiment, loan officers from Rural Credit Cooperatives (RCCs) in Shandong, China were recruited to evaluate randomly selected loan applications and make lending decisions. All the loan files were previously approved with known performance and repayment status. Each loan officer was randomly assigned to one of two incentive groups. One was analogous to pure Personal Responsibility System (PRS), which provided bonuses to loan officers for approved loans that were in performance and imposed penalties on non-performing loans (NPLs).

And the other was PRS with additional penalties for Type II error (i.e. rejecting loans that would have been good). The two groups were further randomized over prior knowledge about probability distribution of the application pool. Results showed that PRS made a risk-averse loan officer inclined to reject loans to avoid risk of penalty. This side effect generated credit rationing, increased Type II error in loan classification and lowered the interest returns of RCCs. Providing prior information about the application pool helped to increase decision accuracy. In theory, this study extended the incentive mechanism design under uncertainty to a behavioral scope. In practical, it contributed to the increase of profitability in financial institutes, alleviated credit rationing and stabilized credit supply in the market.

Three studies in the thesis each targets on one aspect of human behavior. The remaining context is organized as follows: Chapter 2, 3, and 4 each presents one individual study as described above. Chapter 2 (Study 1) deals with non-standard belief. Using overconfidence and self-attribution, study 1 investigates how consumers react to temporary quality failure and perceive food safety risk depending on previous consumption experience. Chapter 3 (Study 2) deals with non-standard decision-making. Using psychological terms such as cognitive dissonance and confirmation bias, study 2 reveals how individual consumers inadequately process information, pay limited attention to signals, and make decisions bias towards their initial choices. Chapter 4 (Study 3) deals with non-standard preference and its market responses. It provides *micro* explanations for *macro* level credit rationing phenomenon in the financial market. Lenders' reference-dependent utility under asymmetric adverse incentive structure broadens literature on institutional mechanism design to a behavioral scope. Finally, Chapter 5 concludes the whole thesis.

References

[1] Camerer, C. and G. Loewenstein. 2004. "Behavioral Economics: Past, Present, Future." in *Advances in Behavioral Economics*, C. Camerer, G. Loewenstein, and M. Rabin, eds., Princeton University Press.

[2] DellaVigna, S. 2009. "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47: 315-372.

[3] Kahneman, D. & Tversky, A. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science*, 185(4157): 1124-1131.

[4] Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47(2): 263-291.

[5] Rabin, M. 1998. "Psychology and Economics." *Journal of Economic Literature*, 36: 11-46.

[6] Rabin, M. 2002. "A Perspective on Psychology and Economics." *European Economic Review*, 46: 657-685.

[7] Stigler G. 1965. "The development of utility theory." Chap. 5 in *Essays in the History of Economics*. Chicago: University of Chicago Press.

2 Chapter 2

Study 1: Consumer Belief and Risk Perception under Temporary Quality Failure

2.1 Introduction

Food safety issues have been a major concern for both public health and food industry. In general, public regulatory agencies and private food companies should be responsible to minimize the food safety risk and provide related information. However, a significant amount of control over these health risks lies in the hands of the consumers, who are the final decision makers of consumption. In this sense, it is of great importance to better understand consumers' risk perception and reaction to related information. Previous studies suggested food safety information is relatively ineffective in changing consumers' behavior (Downs, Loewenstein and Wisdom, 2009), but very few of them offered a reason. This study uses experimental evidence to identify cognitive dissonance and its subsequent behavioral impacts on consumers' risk attitudes and response to information in a food safety context. The results of the experiment provide some explanations for why typical consumers are less responsive to food safety scares.

Cognitive dissonance is a state of discomfort caused by individual holding two contradictory beliefs (Leon Festinger, 1957). By cognitive dissonance theory, past behavior and experience may induce consumers to adjust their beliefs to rationalize their behavior (Akerlof and Dickens, 1982). Wessells, Kline and Anderson (1996) used survey data and showed consumers' perceptions of seafood safety were influenced by their past experiences. Further, the perceptions influenced the anticipated changes in consumption under different information settings. However, in reality, past behavior and experience could also impact consumers' reaction to information, especially when the signals are ambiguous. Confirmatory bias is a natural tendency to reduce dissonance (Frey, 1986). For one thing, it can lead individual to selectively seek confirming evidence and neglect disconfirming evidence. For another, it causes consumers to interpret the ambiguous evidence in a more favorable way (Rabin and Schrag, 1999). Lin, Lee and Yen (2004) found evidence for the first type of confirmation bias, arguing that dietary intakes affect consumers' search of nutrient information.

Our study will extend the existing research in the following ways: First, we differentiate past

experience and one-time shot behavior to investigate the short-term behavioral effect on perception and its interaction with long-term experience. Second, we exgogenously introduce ambiguous signal of the food and investigate how does risk perception change with the signal and past experience. Third, different interpretations of the ambiguous signal would provide evidence for the second type of confirmation bias.

The experiment used in this study was conducted in Ithaca, New York 2007. 61 subjects were invited to a buffet with chicken entree from a local store and 5 other side dishes. 29 in the control group ate the normal main dish and 32 in the treatment group ate chicken mixed with fish sauce. Fish sauce gave a strong smell which performed as an ambiguous external signal indicating the food being potentially tainted. Subjects were further divided by their self-reported past experience with the local store. The results showed risk perceptions were influenced by the existence of the signal, the amount of food eaten and past experience. When food was mixed with fish sauce, users perceived higher risk than non-users conditional on the amount of chicken eaten. However, risk perception decreased as the users ate more chicken, suggesting cognitive dissonance. Further, the users' intention of interpreting the taste signal as being safe was also higher with more chicken eaten, which is consistent with the hypothesis of confirmatory bias. Other responses such as satisfaction of the food and anticipated future consumption were also investigated under different conditions.

Since food safety issues involve a lot of uncertainty, consumers' perceptions depend heavily on their experience and past behaviors. With ambiguous signals, the interpretations are also different depending on the initial perceptions. In order to devise effective communication strategies, we suggest information providers, either policy-makers or private companies, differentiate case by case when offering messages to the public.

2.2 Literature Review

2.2.1 Cognitive Dissonance

Cognitive dissonance theory was originally formulated in the mid-1950s by Leon Festinger and its first complete version was presented in 1957. It is used to refer to the uncomfortable feeling aroused from holding two contradicting attitudes, beliefs or behaviors. To put it more formally, as theorized Festinger, when an individual holds two or more elements of knowledge that are relevant

to each other but inconsistent with one another, a state of discomfort is created. A person who has just bought a car but later finds that the maintenance fee could be very high will feel dissonance because of their former belief that the car is a good choice. When one holds a certain belief but is also forced to act against that, disagreement or dissonance exists between the action and the previously held belief. Motivated by the unpleasant state of dissonance people will further engage in some "psychological work" so as to reduce the inconsistency, and typically this work will support the cognition which is most resistant to change. When returning a car to the dealer is impossible, the owner tends to believe only good car requires higher maintenance fee and maybe the cost is acceptable. A person who just did something irrevocable but is opposite to his long-held belief may regard that belief as less important than before.

In general, to reduce the dissonance, individual could add consonant cognitions, subtract dissonant cognitions, increase the importance of consonant cognitions or decrease the importance of dissonant cognitions. One of the most often assessed ways of reducing dissonance is change in attitudes. Attitude change is expected to be in the direction of the cognition that is most resistant to change. In test of the theory, it is often assumed that the knowledge about recent behavior is usually most resistant to change, since if a person behaved in a certain way, it is often very difficult to undo the behavior. Thus, attitude change would be made consistent with the recent behavior.

Cognitive dissonance theory dominated social psychology from the 1950s until the 1970s. It revolutionized thinking about psychological processes, particularly regarding how actions and outcomes affect attitudes or how behavior and motivation affect perception and cognition. The most influential and widely cited classic experiments are "the post-decision dissonance" (Brehm, 1956), "the induced/forced compliance" (Festinger and Carlsmith, 1959; Freedman, 1965) and "the effort justification" (Aronson and Mills, 1959).

Brehm (1956) examined dissonance theory in post-decision process. In the experiment, participants were asked to make either an easy or a difficult choice between two alternatives. The easy choice was one in which one alternative was much more attractive than the other, whereas the difficult choice was one in which the alternatives are very close in attractiveness. Participants were also asked to evaluate the decision option both before and after the choice decision. According to the theory, after a decision, all of the cognitions that favor the chosen item were consonance, while all the cognitions that favor the rejected item were dissonance. Dissonance could be reduced by viewing

the chosen alternative as more attractive and/or viewing the rejected alternative as less attractive. Brehm found that after people made a difficult decision, they changed their attitudes to become more negative toward the rejected alternative, whereas after an easy decision, their attitudes were not changed.

Festinger and Carlsmith (1959) hypothesized that dissonance should be aroused when a person acts in a way that is contrary to his or her attitudes. To test this, participants were asked to perform a boring task and then were paid either $1 (low justification) or $20 (high justification) to tell another participant that the task was interesting. Since $20 provided sufficient justification for the counter attitude behavior, according to the theory, lying for a payment of $1 should arouse more dissonance than for $20. As expected, participants in low justification group ($1) changed their attitudes to become more positive toward the task; however, those in high justification group ($20) did not change their attitudes and rated the activity boring as before.

Aronson and Mills (1959) designed the first experiment to test the effort justification idea. The idea said dissonance was aroused whenever a person engaged in an unpleasant activity to obtain some desirable outcome. The greater the unpleasant effort required to obtain the outcome, the greater the dissonance. Dissonance could be reduced by exaggerating the desirability of the outcome, since this would add consonant cognitions. In their experiment, women participants needed to undergo a severe or mild "initiation" to become a group member. In the severe initiation condition, women engaged in an embarrassing activity to join the group. In the mild condition however, women engaged only in a simple activity that is not very embarrassing. The group was made dull and boring. But women in the severe condition evaluated the group much more favorably than those in the mild condition, which supported the effort justification idea.

For more innovative experiments on cognitive dissonance, Aronson (1969) and Nisbett and Ross (1999) provide good and comprehensive summaries. Besides these, since late 20th century, there has been renewed interest in cognitive dissonance theory (Beauvois & Joule, 1996; Harmon-Jones & Mills, 1999) and implicit influences on many other contemporary theories (Aronson, 1992). More details regarding the origin and development, challenge and revision of cognitive dissonance theory over the past 50 years could be found in Harmon-Jones and Harmon-Jones (2007).

2.2.2 Confirmation Bias

Another motivational process that was found in line with cognitive dissonance is called confirmatory bias. It is an error in information processing and belief update procedure, which refers to a tendency of selectively collecting information to reinforce the initial belief. As testing of the behavior, dissonance research using a selective exposure paradigm has demonstrated that people are more willing to examine materials that confirm their beliefs than materials that dispute their beliefs (Brock and Balloun, 1967; Frey, 1986). Research using a belief disconfirmation paradigm has shown that when people are exposed to information that challenges their beliefs, they often strengthen their original beliefs (Batson, 1975; Burris, Harmon-Jones and Tarpley, 1997).

Rabin and Schrag (1999) summarize 3 different information-processing problems that will lead to confirmation bias. First, confirmatory bias arises when people have to interpret ambiguous evidence (Keren (1987) and Griffin and Tversky (1992)). A series of experiments in psychology reveal that people tend to misread evidence as additional support for initial beliefs. When facing the same ambiguous information, people with different initial beliefs can move their beliefs even farther apart. Lord, Ross and Lepper (1979) tested this polarization phenomenon. In the experiment, subjects were divided into 2 groups based on their earlier attitudes on death penalty and its deterrent effect. After reading a few randomly selected studies and criticisms on this topic, the subjects were asked to rate their change of attitude. Results showed that the proponents reported they were more in favor of the penalty and its deterrent effect, whereas the opponents reported they were even less in favor of the punishment and the efficacy. Lord, Ross and Lepper explained this as a biased assimilation process which may include "a propensity to remember the strengths of confirming evidence but the weaknesses of the disconfirming evidence, to judge confirming evidence as relevant and reliable but disconfirming evidence as irrelevant and unreliable, and to accept confirming evidence while scrutinizing disconfirming evidence". However, Lord-Ross-Lepper experiment permits an alternative explanation: Since the two groups of people may be predisposed to interpret ambiguous evidence differently, the polarization that proves the difference in interpretation appears to be less relevant to the current beliefs, and thus does not reflect confirmation bias. Darley and Gross (1983) provides a similar but more direct test of polarization based on differing beliefs induced by two ex ante identical groups, and excludes this alternative explanation.

A second situation that may result in confirmatory bias occurs when people must interpret statistical evidence to assess the correlation between phenomena that are separated by time (i.e. hyperactivity and sugar intake, arthritis pain and weather change). Research suggests that inability to identify such correlation is one of the most robust shortcomings in human reasoning (Nisbett and Ross (1980)). Illusory correlation may play an important role in confirmatory bias. People either underestimate the true correlation when they do not perceive it or overestimate some imaginary correlation when they think it is true (Jennings, Amadibile and Ross (1982)).

Third, confirmatory bias can result from people selectively collecting or scrutinizing evidence. A simpler version of selection bias is provided in Wason (1968). In the study, subjects were shown 4 cards with "E", "4", "K" and "7" on each card, and told that each card has a number on one side and a letter on the other. The subjects were then asked which card should be turned over in order to test the hypothesis that "every card with a vowel on one side has an even number on the other". Most subjects chose "E" and "4". While choosing "E" could provide either confirming or disconfirming results depending on whether the number on the other side is even or odd, turning "4" could only provide confirming information if one finds a vowel and no information to test the hypothesis if a consonant is found. In contrast, nearly nobody chose "7". However, turning "7" could disprove the hypothesis if a vowel is found. This is an illustration of individual's willingness to select confirming evidence and to shrink away from information that might disprove the prior hypothesis. A more severe bias could happen when people experience hypothesis-based filtering, in which case, people digest information according to their prior hypotheses and further use the consequent "filtered" evidence as additional support for these hypotheses (Einhorn and Hogarth (1978), Klayman and Ha (1987), Beattie and Baron (1988), Devine, Hirt and Gehrke (1990), Hodgins and Zuckerman (1993) and Zuckerman, Knee, Hodgins and Miyake (1995)). A trader who gets an unclear report of a stock may try to understand it based on his previous impression about the stock. But he will fall into the pitfall if he in turn uses the conclusion he derives from there as further evidence for his investment decision.

Confirmatory bias was widely found in professional fields. Oskamp (1965) found that when clinical psychologists tried to make decisions, their predictive accuracy reached a ceiling in some early point in the information-gathering process. However, confidence about their decisions continued to climb steadily as more information was obtained. Darley and Gross (1983) demonstrated teachers

misread performance of pupils as supporting their initial impressions. Frank and Gilovich (1988) found referees gave significantly more penalty to black-uniformed teams due to the impression that black looked more aggressive. In business management, managers tend to persist with unsuccessful policies (Staw (1976)) and CEOs are overconfident in their acquisition decisions (Bogan and Just (2006)). In finance, traders biases towards early investment (empirical review, see Shleifer (2000)).

2.2.3 Consumer Behavior

Psychological biases such as cognitive dissonance and confirmatory bias have also been extensively applied to consumer behavior. Several articles have provided critical reviews of the theories and have described how the theories are related to consumer behavior (Kassarjian and Cohen, 1965; Cummings and Venkatesan, 1976; Harmon-Jones and Harmon-Jones, 2007). Empirical research generally falls into two categories: (1) effects of dissonance on attitude change and tendency to repurchase, and (2) effects of dissonance on selective information seeking by consumers. In general, studies which have examined the effects of dissonance on attitude change and tendency to repurchase have supported the predictions from the theory (Doob, Carlsmith, Freedman, Landauer and Tom, 1969; Kassarjian and Cohen, 1965). By the foot-in-the-door technique, consumers who care about consistency can make big commitment following a smaller one (Freedman and Fraser (1966) and Pliner, Hart, Kohl and Saari (1974)). However, empirical findings have not supported either a general preference for supportive over non-supportive information or a greater information seeking/avoidance tendency by high dissonance subjects (Freedman and Sears, 1965; Ehrlich et al, 1957; Engel, 1963; LoSciuto and Perloff, 1967). In the marketing situation, it cannot be concluded up to this point that dissonance is relevant to post-decision information seeking. Recently, literature in food safety and public health fill this gap to some extent. Wessells, Kline and Anderson (1996) used survey data and shows consumers' perceptions of seafood safety are influenced by their past experiences. Further, the perceptions influence the anticipated changes in consumption under different hypothetical information concerning seafood. Lin, Lee and Yen (2004) found in field that search for fat and cholesterol information on food labels is less likely among individuals who consume more of these nutrients and thus supports the selective information avoidance tendency that has not been justified in marketing literature.

Research that will further contribute the field includes but not limits to the following directions.

The first direction involves identification of different types of confirmatory bias mentioned above and its impact on consumer risk perception and future behavior. In this study, the authors will demonstrate the contingent existence of cognitive dissonance and individual's tendency of interpreting ambiguous signals to confirm current beliefs. This study will add some new insights concerning consumer behavior in general and food safety framework in particular.

Second, few studies to date have examined the conditions under which dissonance will and will not work. In another experimental setting, we differentiated consumers' responses based on their familiarities to the food they were dealing with and identified the condition under which dissonance would occur (Cao and Just, 2009). This offered some explanations for why consumers were less responsive to public information.

Moreover, most studies in the marketing field adopted the free-choice paradigm and argued it was less possible to use forced compliance paradigm since consumers would not comply with requests of buying sub-optimal goods whenever the best alternative was available. However, in the food consumption situation, we could manipulate this by assigning participants to some certain food, induce dissonance and investigate subsequent behaviors later on. This is also an innovative part of the study compared with other existing research.

Finally, individual decision model that incorporates behavioral anomalies could be further refined to better understand the preferences and utility gains. Many studies have offered fundamentals for this. Rabin and Schrag (1999) models how individual interpret ambiguous information in favor of current belief. Koszegi (1999) used dynamic model to capture the stopping rule in the information seeking process. Yariv (2002) proposes a model that characterizes the taste of consistency, stickiness in behavior and preference over signals. Moreover, in a normative perspective, mechanism designs between behavioral decision maker, strategic information holder and rational social planner could improve the equilibrium in the market interactive context. As of food safety issue, this could rationalize labeling regulation and many other policy interventions regarding public health.

2.3 Experimental Design

This food choice experiment was conducted in fall 2007 at Cornell University (Ithaca, NY). The experiment was designed to measure the actual response of participants to ambiguous signals of food. The weight of food consumed by each participant was recorded and the feeling, expectation and

perception of the food were asked afterwards. We focused on the interaction between participants' past eating behaviors, current risk perceptions and future information processing procedures. Even though the common observation suggests that individuals overestimate the probability of rare events (i.e. Kahneman and Tversky, 1979), research about food choices also reports that consumers tend to underestimate the risk of food-borne illnesses (Hayes et al., 1995). This experimental methodology and the results provide us with an explanation for the underestimation of risk.

In this lunchtime experiment, participants were placed in a situation where they were required to choose whether or not to eat chicken that might be tainted. We considered 2 by 2 conditions with a total of four different cases. First, participants were randomly assigned to two treatment groups. In one group, we mixed the chicken with fish sauce. The fish sauce gave a very strong smell and served as some ambiguous signal that the food was potentially tainted. In the other group, the participants were offered the normal chicken dish. In both treatment groups, chicken was delivered by a popular local processor, called Ithaca Wings. Further, within each group, participants were categorized by users and non-users, depending on whether they had previous experience of eating the food from the local processor.

We hypothesized generally that while individuals would reduce their consumption when facing some strange signals indicating the food being potentially contaminated, they would not cease to consume altogether. This hypothesis is in line with most of the food psychology literature which suggested that individuals have a very hard time resisting food that is immediately available (Boon et al, 1998; Cornell, Rodin and Weingarten, 1989) or that has already been purchased and currently in present within the household (Chandon and Wansink, 2002). In other word, a taste of consistency may cause stickiness in behavior when individuals are making food choice decisions. Further, consumers may be different in sensitivity to the signal due to their previous eating experiences. For this reason, we hypothesized users would sense the strangeness of the food with a higher probability than non-users in the first place. However, the sensitivity would decrease as the participants eat more chicken. This hypothesis is driven from cognitive consistency and confirmatory bias that have been widely found in fields. Subjects would first update their beliefs to be consistent with their previous behaviors and then selectively collect and interpret signals as supporting evidence for their beliefs.

Participants in the experiment were recruited for a "food marketing study", and promised $5

and a meal for their participation. Each session took place at 12:30pm on a Tuesday, Wednesday or Thursday. Subjects were directly informed that the experiment would be conducted by a food psychologist who is a member of the applied economics faculty and that the experiments were not associated with the experimental economics laboratory. Participants entered a room featuring a buffet line to their left and a set of three long tables (seating up to 20 persons each) arranged in a "U" shape on their right. Participants were asked to enter the buffet line and select as much as they liked of each of the foods: boneless fried chicken tenders, French fries, pudding, apple sauce, celery, macaroni salad, soda and bottled water. Subjects were instructed to take at least a little of each item, and each item was to be placed on a separate small plate on their tray. At the end of the buffet line, all plates were weighted individually, and participants were then told to be seated at one of the three long tables on the outer edge of the "U" and to wait for instruction from the experimenter before beginning to eat. By sitting on the outside of the "U", each individual could easily see and hear all of the other participants. After the experimenter checked each tray to make sure everyone had complied with instructions, the subjects began eating. After completing their meals, the participants' plate were again weighted to determine how much of each item had been consumed. Each subject was then asked to respond to a survey. Following completing the survey, participants were asked to discuss their experiment with the experimenter and provide their impressions.

2.4 Experimental Results

2.4.1 Summary Statistics and Treatment Groups

A total of 61 participants completed the experiment. Summary statistics could be found in Table 1. As shown in the table, we had a good control over age and the body mass index (BMI), however, the two treatment groups are slightly different in gender, height and weight, with fish sauce group having more female and thus, lower height and weight measures than the non fish sauce group. In later analysis, we controlled these factors and the further robustness check ensured that our main results do not change with these issues.

Each group was further decomposed based on participants' responses to the question "how many times did you eat at Wings (the local processor)". Those who reported non-zero visiting time were

Table 1: Study 1_Summay Statistics by Treatment

Variable	All	fishsauce	non-fishsauce	P-value
Gender (Female=2)	1.525	1.667	1.357	0.015
	(0.50)	0.48	0.49	
Age	19.721	19.727	19.714	0.977
	(1.69)	1.42	2.00	
Height	67.697	66.682	68.893	0.008
	(3.31)	3.10	3.19	
Weight (lbs.)	145.377	139.879	151.857	0.080
	(26.65)	22.42	30.03	
BMI	22.166	22.020	22.338	0.648
	(2.68)	2.37	3.04	
No. obs.	61	28	33	

called (experienced) users. And others who had no previous experience eating at Wings were called (inexperienced) non-users. 15 out of 28 participants in the fish sauce group were categorized as users and 17 out of 33 in the non fish sauce group were users. Since we focused on the interaction between behavior, perception and information processing, this way of organizing participants enabled us the most flexibility to test our hypotheses. While fish sauce served as an ambiguous signal of the food being potentially tainted, we still expected different sensitivities and responses to the signals due to different past experiences.

Several questions were asked as manipulation check for the fish sauce treatment and the division of user and non-user. In response to the question "I ate more chicken than usual", fish sauce group reported a significant decrease (an average difference of 1.8 out of 9-point scale, F=13.64, P=0.00). For "the chicken tastes better than usual", participants in non fish sauce group reported higher rates (F=8.64, P=0.00). More participants in fish sauce group agreed with the statement "the chicken didn't' taste quite right" (F=9.19, P=0.00). For all the questions above, there was no significant difference between users and non-users, which implied our division of group is independent between each other. In order to check the correctness of users and non-users, we further asked questions such as "how many times did you eat carryout from Wings", "how many times did you eat last year at Wings" and "when is the last time you ate at Wings" etc. All the results were consistent with the ones from the original question "how many times did you eat at Wings". Once again, checking the responses to these questions, there was no significant difference due to the fish sauce treatment and the independence was verified.

2.4.2 Behavior

In order to investigate the impact of past behavior on perception and dissonance feeling, variables were organized as different stages in the cognitive procedure. We started with a group of measures called "Behavior", which included the amount of food taken, remaining and eaten and the percentage of the food eaten by each individual. The amount of food eaten was then used as control over the intensity of the behavioral impact, with more food eaten representing stronger behavior. The second group of variables "Perception" measured the general feelings, such as how tasty the chicken was, how high the quality was, did the chicken taste right or tainted and the expectation of the last piece eaten, etc. The third group "Dissonance" further linked the feelings with potential food safety risk (i.e. birdflu), and measured how individual felt about the risk and their judgments on the related issues. The last group "Future Behavior" elicited the participants' willingness to pay (WTP) based on their past eating behaviors and their influenced risk perceptions.

Table 2 gave the food choice behavior between groups. On average, people took 150.35 grams of chicken and ate 127.85 grams. For the fish sauce effect, in general, people ate less when the chicken was mixed with fish sauce (F=8.16, P=0.00). This trend could also be found separately in both user group (125.5 vs. 154.75) and the non-user group (87.82 vs. 154.2, F=7.64, P=0.01), even though the difference in the user group is not significant. This finding was consistent with our hypothesis that when the food was a little bit strange, consumers would decrease their consumption, but would not stop eating it altogether. For the user effect, users ate slightly more than non users (138.04 vs. 118.94). Comparing choices between groups, in the fish sauce group, users took and ate significantly more chicken than the non-users (F=3.48, P=0.07), but in the non fish sauce group, users and non-users were not different from each other significantly (F=0.00, P=0.98). This phenomenon could be explained in this way: Experience did not play much role when the food was normal (as in the non fish sauce group); But as the condition of the food changed (as in the fish sauce group), users had higher tolerance to the conditions of the food due to their past experiences.

Since chicken eaten may also be influenced by risk perceptions and feelings in the first place, we need to find instrument variables for eating behavior, predict it and make the predicted value being independent of previous perceptions and feelings. There are three groups of variables that can serve as instruments: First, exogenous treatments, i.e. fish sauce and user indicators are determined

Table 2: Study 1_Food Choice Behavior by Group

Variable	All	(fs=0, user=0)	(fs=0, user=1)	(fs=1, user=0)	(fs=1, user=1)	t(user)_fs	t(user)_nfs	t(fs)_user	t(fs)_nuser
	col0	col1	con2	con3	col4	col4-col3	col2-col1	col4-col2	col3-col1
ChickenTaken	150.350	174.400	165.750	113.118	155.813	1.87	-0.29	-0.36	-2.46
	(73.35)	(84.45)	(68.71)	(49.65)	(77.74)				
Remaining	22.131	20.200	10.154	25.294	30.313	0.35	-1.00	1.62	0.41
	(35.26)	(34.18)	(17.04)	(35.19)	(46.04)				
Eaten	127.850	154.200	154.750	87.824	125.500	1.87	0.02	-1.25	-2.68
	(69.09)	(78.38)	(63.94)	(58.68)	(57.22)				
% Eaten	0.826	0.900	0.943	0.730	0.770	0.35	0.83	-2.20	-1.70
	(0.27)	(0.17)	(0.10)	(0.37)	(0.29)				
No. obs.	61	15	13	17	16				

independent of individual psychological feelings, but will influence eating behaviors. Second, the amount of food taken at the beginning of the experiment, which indicate subjects' hungry level, are considered independent of emotional feelings, as well. In addition, other types of food available in the buffet line can be used as substitutes of the main dish - chicken. Finally, demographics such as age, gender, BMI can have impacts on eating behavior but not on risk feelings.

Table 3 reported regression results about eating behavior on 4 different combinations of instruments. In general, adding fish sauce to the chicken entree decreased eating amount by roughly 20 grams. Whether a subject had previous consumption experience or not (User or Non-user) does not significantly impact eating amount. Subjects ate about 78% of chicken they took from the buffet line. However, the amount of side dishes eaten, i.e. French fries, pudding, macaroni salad, celery and apply sauce, do not have any significant impact on chicken eaten. Further, effect of demographic variables are not significant either.

In the following context, use the regression in the first column of Table 3 to predict chicken eaten, and use the predicted value (which is now independent of risk perceptions) as explanatory variable to investigate effects of eating behavior on perceptions and feelings. In the tables that follows (i.e. Table-4 to 7), variable *Chickeaten* is the predicted value.

2.4.3 Perception

Table 4 listed a few measures of perceptions and their changing patterns with the eating behaviors. Each column represented a perception measure, with individuals' agreement in a 9-point scale. Regressing each of the perception measures on the treatment dummies fish sauce and user, and be-

Table 3: Study 1_Instrument for Behavior

	1	2	3	4
VARIABLES	chickeaten	chickeaten	chickeaten	chickeaten
fishsauce	-23.75**	-22.41**	-20.93**	-19.92**
	(-2.384)	(-2.386)	(-2.329)	(-2.295)
user	7.148	8.03	7.424	9.006
	(0.76)	(0.92)	(0.83)	(1.06)
chicktaken	0.781***	0.762***	0.809***	0.782***
	(10.93)	(11.79)	(12.42)	(13.00)
ffeaten	0.133	0.14		
	(1.21)	(1.54)		
puddingeate	0.0339	0.0307		
	(0.40)	(0.37)		
mseaten	0.0575	0.0843		
	(0.81)	(1.26)		
celeaten	-0.0416	-0.0518		
	(-0.380)	(-0.494)		
aseaten	-0.0301	-0.0227		
	(-0.292)	(-0.243)		
height	-1.547		-1.736	
	(-0.814)		(-0.943)	
bmi	2.241		2.325	
	(1.18)		(1.34)	
age	2.068		4.131	
	(0.68)		(1.61)	
male	2.471		6.076	
	(0.19)		(0.49)	
Constant	21.45	7.691	-4.163	16.68
	(0.16)	(0.58)	(-0.0327)	(1.38)
Observatior	60	60	60	60
R-squared	0.821	0.813	0.81	0.793

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

havior variables chicken eaten, in addition to the interaction terms between fish sauce and chicken eaten (*FsChick*), and user and chicken eaten (*UseChick*), table-3 recorded all the estimated parameters for each perception. Based on the results, participants could sense the fish sauce treatment correctly in the first place, but their perceptions were influenced in an opposite way as they ate more chicken. For example, for the statement "the chicken is very tasty" (*Chicken2*), people in fish sauce group reported 2.78 points lower than those in the non fish sauce group, but as they ate more chicken, their agreement to this statement was increasing with an additional slope of 0.0157. Similarly, participants behaved the same trend for other statement like "the chicken is of high quality" (*Chicken3*), "the chicken is better than typical" (*Chicken4*) and "the expectation of the last piece of chicken you ate" (*Expect2*), etc. In contrary, for some statement regarding the negative perception of the food, such as "the chicken doesn't taste quite right" (*Chicken5*) and "the chicken tastes tainted" (*Chicken6*), participants in fish sauce group on average reported a higher agreement in the constant term (3.34 and 2.32 respectively), but an additional lower slope (-0.0156 and -0.0114) implying that the more they ate, the less negatively they rated the food. These findings suggested that people did

Table 4: Study 1_Impact of Food Choice on Perception

VARIABLES	Chicken1	Chicken2	Chicken3	Chicken4	Chicken5	Chicken6	Expect2
Fishsauce	-0.1	-2.777***	-2.750***	-1.962*	3.343***	2.317*	-2.323*
	(-0.0849)	(-2.808)	(-2.715)	(-1.744)	(2.74)	(2.00)	(-1.997)
User	0.00675	1.34	1.005	-0.528	0.414	0.18	-0.722
	(0.01)	(1.30)	(1.01)	(-0.478)	(0.33)	(0.15)	(-0.633)
ChickenEaten	0.0105*	-0.00183	-0.00262	-0.00132	0.00461	0.00173	-0.00847
	(1.71)	(-0.363)	(-0.487)	(-0.225)	(0.74)	(0.29)	(-1.414)
FsChick	0.0055	0.0157**	0.0159**	0.0111	-0.0156*	-0.0114	0.0136*
	(0.67)	(2.28)	(2.26)	(1.42)	(-1.845)	(-1.424)	(1.70)
UserChick	-0.00127	-0.00521	-0.00522	0.00438	-0.00428	-0.00296	0.00604
	(-0.161)	(-0.755)	(-0.761)	(0.58)	(-0.504)	(-0.367)	(0.78)
Constant	4.206***	6.712***	6.077***	5.433***	2.435**	2.597**	5.621***
	(4.14)	(8.02)	(6.94)	(5.60)	(2.36)	(2.65)	(5.66)
Obs.	60	59	59	60	59	59	59

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

feel cognitive dissonance and confirmatory bias. On the one hand, eating chicken with fish sauce made them feel that the food was somewhat strange. This raised a dissonance feeling in their mind as oppose to their cognition that they always ate something safe. On the other hand, individuals would have a tendency to reduce the dissonance feeling. The more they ate the more dissonance they felt, the stronger the tendency was. Since people could not change their past eating behavior, the only way feasible for them to reduce the dissonance was to perceive the food in a more favorable way. And this is just what we found in the data.

2.4.4 Dissonance

A further investigation of the dissonance measures showed that even though people perceived some objective characteristics of the food in a right way (Table 4), they still performed biases when making judgment involving risk. In table 5, the (hypothetical) statement "the chicken is partially infected with bird flu" (*Birdflu6*) was tested. Similar as before, users still reported a higher rate (1.41) than non users, implying that previous experience enabled them to be more sensitive. But once again, the amount of the food eaten had an opposite effect on the judgment. For each 1 more gram of chicken they ate, the users lowered their level of agreement by 0.008. Given the fact that on average, people ate more than 150 grams, this effect was large enough to offset their original sensitivity (the 1.41 points higher in the constant term). Further, fish sauce treatment had an additional negative effect of 1.43, which means that in fish sauce group, people were even more

Table 5: Study 1_Impact of Food Choice on Dissonance

VARIABLES	All Birdflu6	All Because9	FS=0 Because9	FS=1 Because9	All Because4	User=0 Because4	User=1 Because4	All Because5	User=0 Because5	User=1 Because5	All Because7	All Because10
FishSauce	-1.426**	1.168			2.029**	1.898*	2.352	0.628	1.024	-0.12	-0.0315	-0.276
	(-2.007)	(1.14)			(2.13)	(1.74)	(1.25)	(0.45)	(0.70)	(-0.0416)	(-0.0999)	(-0.537)
User	1.409**	2.010*	2.500*	0.764	-0.596			2.351				
	(2.02)	(1.87)	(1.90)	(0.45)	(-0.599)			(1.58)				
ChickenEaten	-0.00359	0.00678	0.00891*	-0.0109	0.0125**	0.0121**	0.0204**	0.0164**	0.0181**	-0.0048	0.00423*	0.00652*
	(-0.970)	(1.29)	(1.95)	(-1.353)	(2.58)	(2.41)	(2.57)	(2.30)	(2.66)	(-0.399)	(1.81)	(1.68)
FsChick	0.00759	-0.0106			-0.00603	-0.00662	-0.00691	-0.00371	-0.0142	0.00713		
	(1.54)	(-1.486)			(-0.913)	(-0.819)	(-0.562)	(-0.383)	(-1.307)	(0.38)		
UserChick	-0.00805*	-0.00961	-0.0158*	0.005	0.00718			-0.0180*				
	(-1.687)	(-1.340)	(-2.020)	(0.38)	(1.08)			(-1.830)				
Constant	2.112***	6.349***	6.226***	7.953***	4.653***	4.793***	3.842***	3.738***	3.816***	6.159***	7.568***	6.594***
	(3.45)	(7.29)	(7.93)	(9.47)	(5.78)	(5.53)	(2.90)	(3.15)	(3.27)	(3.08)	(17.88)	(9.57)
Obs.	60	59	27	32	59	32	27	58	32	26	59	58

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

reluctant to admit the food was partially infected.

Moreover, the amount of food eaten also had significant positive impact on the judgment such as "I didn't believe it (bird flu) would hurt me" (*Because7*) and "the food was safe" (*Because10*). The statements "I ate chicken because I am hungry" (*Because4*) and "I usually eat what's in front of me" (*Because5*) could be used as arguments to justify previous eating behaviors to some extent. And for these, the amount of food eaten also had the same effect. The more participants ate, the higher the desire they wanted to justify. Regarding "I ate chicken because it was a study" (*Because9*), even though one could observe significant effects, the interpretation could be a bit ambiguous.

2.4.5 Future Behavior

In addition to the perceptions and dissonance feelings, variables regarding future behaviors were also collected. First, participants were asked about their WTP for the whole meal (*Limit*). Table 6 shows the results. While fish sauce decreased the users' WTP by $1.88 (close to be significant, t=-1.45), it did not have strong effect for non-users. The amount of food eaten increased the WTP of the users, but once again had no effect on non-users. Comparing the fish sauce treatment effect and the effect of eating amount (marginal effect on the slope term), one could observe that among users, the slope effect could overweight the treatment effect, given that on average, people ate 150 grams chicken. Though users can sense the strangeness about the food, they lose their clue as eating more.

The participants were also given a chance to trade their $5 payment to a larger amount in gift certificates redeemable for food at Wings over Ithaca, the local processor which offered the food in

Table 6: Study 1_Impact of Food Choice on Future Behavior

VARIABLES	All Limit	User=1 Limit	User=0 Limit	FS=1 Limit	FS=0 Limit	All Trade1	All Trade0	User=1 Trade0	User=0 Trade0	If Trade Trade1	If Trade Trade1
FishSauce	−0.308	−1.878	1.281			−0.478					
	(−0.253)	(−1.450)	(0.64)			(−0.257)					
User	−1.588			0.0213	−1.234	3.792				−2.517*	−2.981**
	(−0.627)			(0.01)	(−0.250)	(0.97)				(−1.832)	(−2.160)
ChickenEaten	0.0143	0.0209*	0.0191	0.0365**	0.00798	0.0335*	0.00172**	0.000703	0.00236**	−0.00848	
	(1.30)	(1.95)	(1.42)	(2.30)	(0.47)	(1.86)	(2.07)	(0.48)	(2.32)	(−0.847)	
UseChick	0.00973			−0.0195	0.0173	−0.0463*					
	(0.56)			(−0.848)	(0.59)	(−1.701)					
Constant	4.479**	4.223**	3.068	2.824*	4.782	5.961**	0.496***	0.617***	0.438***	14.52***	13.41***
	(2.47)	(2.21)	(1.25)	(1.71)	(1.62)	(2.11)	(4.11)	(2.78)	(3.07)	(8.99)	(13.91)
Obs.	60	28	32	33	27	59	60	28	32	42	43

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

the experiment. The procedure was that participants first chose whether to trade or not (*Trade0*). If yes, they were then asked to pick an integer number that they would be willing to trade ranging from \$5 to \$20 (*Trade1*). Then a 16-sided die with sides numbered 5 to 20 was rolled. If the roll was greater than the amount the participant picked, he/she would be given the amount of the roll in gift certificate. If the roll is less than his/her willingness to accept, he/she would still keep the \$5 in cash. General hypothesis would be people chose to trade if they liked the food. Given they chose to trade, people would be willing to claim for lower amount in gift certificate if they liked the food more. Moreover, by the design of the game, people would make trade-off between the amount they were willing to accept and the possibility they could win. If a participant valued \$1 certificate to be the same as \$1 in cash, he/she should claimed for \$9, which was the optimal choice considering the trade-off. However, in the experiment, on average, users claimed for about \$10.5 and non-users claimed for \$13.4, both higher than \$9. This implied participants valued \$1 in gift certificate less than \$1 in cash, which was normal.

44 out of 61 participants chose to trade, among whom 23 were users (with 11 in fish sauce group and 12 in non fish sauce group) and 21 were non users (with 11 in fish sauce group and 10 in non fish sauce group). Regression of willingness amount to trade (*Trade*1) in Table 6 suggested non-users were more possible to be impacted by the eating behaviors. The more they ate, the higher the amount their willingness to accept was. Users' choices didn't bear the same effect from eating behaviors. (Column under "*All, Trade1*") Tobit model with data censored between 5 and 20 gave the same results as OLS regression. Further, the amount of chicken eaten positively impacted the possibility to trade (*Trade0*) among non-users (Column "*User=0, Trade0*"). As the participants

ate more chicken, it was more likely for them to choose to trade the $5 cash for a larger amount of gift certificate, implying a taste of consistency among consumers, especially when they were new customers. Further, given participants chose to trade, non-users tended to claim almost $3 higher than users (Columns "*If Trade, Trade*1"). Lack of previous experiences among non-users could be an explanation for this gap.

2.5 Discussion

This study suggested consumers had a taste of consistency in choosing what to eat and would be subject to confirmatory bias when making judgments. When facing an ambiguous signal about the food quality, consumers were more likely to neglect the signal and rate the food as more favorable if they ate more. Even though consumers who had previous experience with the food would notice the signal with a higher probability, their judgments were also mitigated by a larger magnitude depending on the amount they ate before. This finding offered an explanation for why consumers were universally irresponsive to the public food safety information. Non-users might overlook the food safety issue due to inexperience. But the users could also misperceive the potential risk so as to justify their previous eating behavior and reduce the dissonance feelings in their mind.

A few caveats need to be pointed here. First, we did have significant difference in gender, height and weight across groups. But the main results did not change when the demographics were included in the model.

Second, as shown in Table 7 47.5% of the non-users and 31.25% of the users in the fish sauce group ate all of the chicken, which were slightly lower than the percentage in the non fish sauce group (54% of the participants ate all the chicken). Of all those participating in the experiment, only 2 out of 33 (6%) in the fish sauce group refused to eat any of the chicken they had taken. On the one hand, people would decrease eating amount when facing ambiguous signals and the amount decreased depend on their previous experiences. On the other hand, very few would cease to eat the chicken altogether. In terms of policy concern, some existing estimation of the changes in consumption due to food safety issues might be exaggerated in the field. Given the behavioral pattern found in this study, substantial efforts would be needed to fully eliminate consumption changes when food is recalled.

For the experimental design, since there was no other meat immediately available, participants

Table 7: Study 1_Participants Consuming None, Some and All of the Chicken by Group

Chicken Consumed	User=0 FishSauce=0	User=0 FishSauce=1	User=1 FishSauce=0	User=1 FishSauce=1
All	8	8	7	5
	(54%)	(47.5%)	(54%)	(31.25%)
Some	7	8	6	10
	(47%)	(47.5%)	(46%)	(62.5%)
None	0	1	0	1
	–	(6%)	–	(6.25%)
Obs.	15	17	13	16

might have eaten some chicken when they otherwise would have avoided. Future work could address this potential interaction with substitutes by adding more choices of main dish in addition to the side dishes (i.e. French Fries and pudding, etc).

Moreover, the food served in the study was delivered by a local restaurant right before consumption, thus, the level of consumption might reflect an inherent trust that a food retailer would not provide tainted food due to liability concerns. Even though without the local food processor, participants' eating behaviors might also to some extent reflect their original perceptions regarding the food. Being unable to tease out those effects would yield potentially biased results. This study chose 4 groups of instrumental variables to address this problem.

References

[1] Akerlof, George A & Dickens, William T, 1982. "The Economic Consequences of Cognitive Dissonance," American Economic Review, 72(3): 307-319.

[2] Aronson, E., "The Theory of Cognitive Dissonance: A Current Perspective", in L. Berkowitz (Ed.), Advances in Experimental Social Psychology, 4 (1969), New York: Academic.

[3] Aronson, E., "The Return of the Repressed: Dissonance Theory Makes a Comeback", Psychological Inquiry, 3 (1992), 303-311.

[4] Aronson, E. and Mills, J., "The effect of severity of initiation on liking for a group", Journal of Abnormal and Social Psychology, 59 (1959), 177–181.

[5] Batson, C.D., "Rational processing or rationalization? The effect of disconfirming information on a stated religious belief", Journal of Personality and Social Psychology, 32 (1975), 176–184.

[6] Beattie, J. and J. Baron, "Confirmation and Matching Bias in Hypothesis Testing", Quarterly Journal of Experimental Psychology: Human Experimental Psychology, XL (1988), 269-297.

[7] Beauvois, J.L., and Joule, R.V., A radical dissonance theory, London: Taylor and Francis, (1996).

[8] Beauvois, J.L., and Joule, R.V., "A Radical Point of View on Dissonance Theory", in E. Harmon-Jones and J. Mills (Ed.), Cognitive Dissonance: Progress on a Pivotal Theory of Social Psychology, Washington, DC: American Psychological Association, (1999).

[9] Bogan, V. and D. Just, "What Drives Merger Decision Making Behavior: Don't Seek, Don't Find and Don't Change Your Mind", Journal of Economic Behavior and Organization, 72(3) (2009), 930-943.

[10] Brock, T.C., and J.C. Balloun , " Behavioral receptivity to dissonant information", Journal of Personality and Social Psychology, 6 (1967), 413–428.

[11] Brehm, J. W., "Post-decision Changes in the Desirability of Alternative", Journal of Abnormal and Social Psychology, 52 (May 1956), 384-389.

[12] Burris, C.T., E. Harmon-Jones and W.R. Tarpley, "By faith alone": Religious agitation and cognitive dissonance", Basic and Applied Social Psychology, 19 (1997), 17–31.

[13] Coob, A. N., J. M. Carlsmith, J. Freedman, T. K. Landauer, and S. Tom, "Effect of Initial Selling Price on Subsequent Sales", Journal of Personality and Social Psychology, 11(1969), 354-350.

[14] Cummings, W. H. and M. Venkatesan, "Cognitive Dissonance and Consumer Behavior: A Review of the Evidence", Journal of Marketing Research, XIII (1976), 303-308.

[15] Darley, J. and P. Gross, "A Hypothesis-Confirming Bias in Labeling Effects", Journal of Personality and Social Psychology, XLIV (1983), 20-33.

[16] Devine, P. G. and E. R. Hirt and E. M. Gehrke, "Diagnostic and Confirmation Strategies in Trait Hypothesis Testing", Journal of Personality and Social Psychology, LVIII (1990), 952-963.

[17] Doob, A. N., J. M. Carlsmith, J Freedman, T. K. Landauer and S. Tom, "Effect of Initial Selling Price on Subsequent Sales", Journal of Personality and Social Psychology, 11 (April 1969), 345-350.

[18] Downs, Julie S., George Loewenstein, and Jessica Wisdom. 2009. "Strategies for Promoting Healthier Food Choices." American Economic Review, 99(2): 159–64.

[19] Ehrlich, D., I. Guttman, P. Schonbach and J. Mills, "Post-decision Exposure to Relevant Information", Journal of Abnormal and Social Psychology, 54 (Jan 1957), 98-102.

[20] Einhorn, H. and R. Hogarth, "Confidence in Judgment: Persistence of the Illusion of Validity", Psychological Review, LXXXV (1978), 395-416.

[21] Engel, J. F., "Are Automobile Purchases Dissonant Consumers?" Journal of Marketing, 27 (April 1963), 55-58.

[22] Festinger, L., A Theory of Cognitive Dissonance, Stanford, CA: Stanford, (1957).

[23] Festinger, L., & J.M. Carlsmith, "Cognitive consequences of forced compliance", Journal of Abnormal and Social Psychology, 58 (1959), 203–210.

[24] Frank, M. G. and T. Gilvich, "The Dark Side of Self and Social Perception: Black Uniforms and Aggression in Professional Sports", Journal of Personality and Social Psychology, 54 (1988), 74-85.

[25] Freedman, J. L. and S. C. Fraser, "Compliance without Pressure: The Foot-In-The-Door Technique", Journal of Personality and Social Psychology, 4 (1966), 195-202.

[26] Freedman, J. L. and D. O. Sears, "Selective Exposure", in L. Berkowitz, Advances in Experimental Social Psychology, Vol. 2 New York: Academic Press, (1965).

[27] Frey, D., "Recent research on selective exposure to information", in L. Berkowitz (Ed.), Advances in experimental social psychology, 19 (1986), 41–80. New York: Academic Press.

[28] Friedrich, James, "Primary error detection and minimization (PEDMIN) strategies in social cognition: A reinterpretation of confirmation bias phenomena", Psychological Review, 100(2) (Apr 1993), 298-319.

[29] Gallimore P., "Confirmation bias in the valuation process: a test for corroborating evidence", Journal of Property Research, 13(4) (Dec. 1996), 261-273.

[30] Griffin, D. and A. Tversky, "The Weighting of Evidence and the Determinants of Confidence", Cognitive Psychology, XXIV (1992), 411-435.

[31] Harmon-Jones, E. and C. Harmon-Jones, "Cognitive Dissonance Theory: After 50 Years of Development", Zeitschriftfur Sozialpsychologie, 38(1 2007), 7-16.

[32] Harmon-Jones, E. and J. Mills, Cognitive Dissonance: Progress on a Pivotal Theory of Social Psychology, Washington, DC: American Psychological Association, (1999).

[33] Hodgins, H. S. and M. Zuckerman, "Beyond Selecting Information: Biases in Spontaneous Questions and Resultant Conclusions", Journal of Experimental Social Psychology, XXIX (1993), 387-407.

[34] Jennings, D. L., M. R. Lepper and L. Ross, "Informal Covariation Assessment: Data-Based versus Theory-Based Judgments", in Judgment under Uncertainty: Heuristics and Biases, D. Kahneman, P. Slovic and A. Tversky (Ed.), Cambridge: Cambridge University Press, (1982), 211-230.

[35] Jonas, E., S. Schulz-Hardt, D. Frey and N. Thelen, "Confirmation Bias in Sequential Information Search after Preliminary Decisions: An Expansion of Dissonance Theoretical Research on Selective Exposure to Information", Journal of Personality and Social Psychology, 80(4) (Apr 2001), 557-571.

[36] Kassajian, H. H. and J. B. Cohen, "Cognitive Dissonance and Consumer Behavior", California Management Review, 8 (Fall 1965), 55-64.

[37] Keren, G., "Facing Uncertainty in the Game Bridge: A Calibration Study", Organizational Behavior and Human Decision Processes, XXXIX (1987), 98-114.

[38] Klayman, J. and Y. Ha, "Confirmation, Disconfirmation and Information in Hypothesis Testing", Psychological Review, XCIV (1987), 211-228.

[39] Koszegi, B., "Self-image and Economic Behavior", MIT dissertation.

[40] Lin, C. J., J. Lee and S. T. Yen, "Do Dietary Intakes Affect Search of Nutrient Information on Food Labels?", Social Science and Medicine, 59 (2004), 1955-1967.

[41] Lord, C. G., L. Ross and M. R. Lepper, "Biased Assimilation and Attitude Polarization: The Effect of Prior Theories on Subsequently Considered Evidence", Journal of Personality and Social Psychology, XXXVII (1979), 2098-2109.

[42] LoSciuto, L. and R. Perloff, "Influence of Product Preference on Dissonance Reduction", Journal of Marketing Research, 4 (March 1967), 286-290.

[43] Mittelstaedt, R., "A Dissonance Approach to Repeat Purchasing Behavior", Journal of Marketing Research, 6 (November 1969), 444-446.

[44] Mynatt, Clifford R., Doherty, Michael E. and Tweney, Ryan D., "Confirmation bias in a simulated research environment: An experimental study of scientific inference." The Quarterly Journal of Experimental Psychology, 29(1) (Feb 1977), 85-95.

[45] Nisbett, R. E.,and L. Ross, Human Inference: Strategies and Shortcomings of Social Judgment, Englewood Cliffs, NJ: Prentice-Hall, (1980).

[46] Nisbett, R. E.,and L. Ross, Person and the Situation – Perspectives of Social Psychology, Philadelphia: Temple University Press, (1991).

[47] Oskamp, S., "Overconfidence in Case-Study Judgments", Journal of Consulting Psychology, 29 (1965), 261-265.

[48] Pliner, P. and H. Hart, J. Kohl and D. Saari, "Compliance without Pressure: Some Further Data on the Foot-In-The-Door Technique", Journal of Experimental Social Psychology, 10 (1974), 17-22.

[49] Piggott, N. and T. Marsh, "Does Food Safety Information Impact U.S. Meat Demand?" American Journal of Agricultural Economics, 86 (2004), 154-174.

[50] Rabin, M. and J. L. Schrag, "First Impressions Matter: A Model of Confirmation Bias", Quarterly Journal of Economics, 114(1) (Feb. 1999), 37-82.

[51] Schlenker, W. and S.B. Villas-Boas, "Consumer and Market Responses to Mad-Cow Disease", CUDARE Working Paper 1023, Department of Agricultural and Resource Economics, University of California-Berkeley, (Nov. 2006).

[52] Shleifer, A., "Inefficient Markets", Clarendon Lectures, Oxford University Press, (2000).

[53] Staw, B. M., "Knee-deep in the Big Muddy: A Study of Escalation Commitment to a Chosen Course of Action", Organizational Behavior and Human Performance, 16 (1994), 857-869.

[54] Wason, P., "Reasoning about a Rule", Quarterly Journal of Experimental Psychology, 20 (1968), 273-281.

[55] Wessells, C. R., J. Kline and J. G. Anderson, "Seafood Safety Perceptions and Their Effects on Anticipated Consumption under Varying Information Treatments", Agricultural and Resource Economics Review, (April 1996), 12-21.

[56] Yariv, L., "I'll See It When I Believe It – A Simple Model of Cognitive Consistency", Cowles Foundation Discussion Paper NO. 1352, Yale University, (Feb. 2002).

[57] Zuckerman, M., C. R. Knee, H. S. Hodgins and K. Miyake, "Hypothesis Confirmation: the Joint Effect of Positive Test Strategy and Acquiescence Response Set", Journal of Personality and Social Psychology, LXVIII (1995), 52-60.

3 Chapter 3

Study 2: Cognitive Dissonance, Confirmation Bias and Inadequate Information Processing

3.1 Introduction

Psychological biases such as cognitive dissonance and confirmatory bias have been extensively applied to consumer behavior. Empirical research generally falls into two categories: (1) effects of dissonance on attitude change and tendency to repurchase, and (2) effects of dissonance on selective information seeking by consumers. In general, studies along the first line have supported the predictions from the theory (Doob, Carlsmith, Freedman, Landauer and Tom, 1969; Kassarjian and Cohen, 1965). However, empirical findings have not supported either a general preference for supportive over non-supportive information or a greater information seeking/avoidance tendency by high dissonance subjects (Freedman and Sears, 1965; Ehrlich et al, 1957; Engel, 1963; LoSciuto and Perloff, 1967).

Recently, literature in food choices and public health fills this gap to some extent. Wessells, Kline and Anderson (1996) uses survey data and shows consumers' perceptions of seafood safety are influenced by their past experiences. Further, the perceptions influence the anticipated changes in consumption under different hypothetical information concerning seafood. Lin, Lee and Yen (2004) finds in field that search for fat and cholesterol information on food labels is less likely among individuals who consume more of these nutrients and thus supports the selective information avoidance tendency.

Some general critiques for the above research are: First, most studies fail to establish an exclusive relationship between their results and the theory. Second, endogeneity problem among risk, perception and behavior may further bias the result (Shogren and Stamland, 2007). With these concerns, an incentive compatible experimental design can have the advantage of isolating alternative competing interpretations and identifying the causal relationship accurately. Relevant designs in the previous literature include two major directions: food safety preference and valuation (such as Hayes et al, 1995) and informational impact on consumer behavior (such as Lusk et al, 2004a).

The objective of this study is to investigate the following research questions: First, how does

individual consumer react to information regarding food safety issue in general? Second, what are the psychological factors that may affect consumers' responses to the information? Third, if the psychological factors exist, to what extent do they affect the decision-making? And fourth, how to mediate the factors that impact consumers' rational decision making so as to make the public communication more effective?

In order to make clear the above questions, a two-group between-subject experiment was conducted. In the study, participants were instructed to bid for chocolate candy bars with 3 different flavors through BDM (Becker-Degroot-Marschak) auction. The auction lasted for 3 rounds. Aflatoxin food safety information was provided to the participants at the beginning of the second and the third round. Of all the 3 flavors, peanut flavor was highly related to the food-borne illness, followed by almond flavor, and plain flavor did not get involved in the food safety issue in the study. For all the 3 rounds, risk perceptions and bids were collected. The only difference between control and treatment group was that participants bid for all 3 flavors and randomly got one flavor to be binding in the control group, whereas participants in the treatment group freely chose one out of the 3 flavor to bid at the very beginning.

Results showed that as soon as participants made their free choice, they bid higher prices for the candy bars. The result was robust when the factors such as risk perception and preference were controlled. This result suggested consumers experienced cognitive dissonance after committing to one product and the higher bids was used as self-compliance device to justify their previous choices.

Further, when participants were exposed to safety information involving the products they chose, though risk perceptions increased, the changing magnitude in the treatment group was much lower than that of control group. As a natural tendency to reduce dissonance feelings, consumers selectively ignored conflicting (risk) information and behaved with confirmatory bias. In terms of market responses, the demand curves were less likely to shift downwards after the risk information shock. While loyal customers were more tolerant to negative news regarding their favorite products, they were also less sensible in picking up crucial information when health or food safety issues were involved.

The remaining paper is organized as follows: Section 3.2 describes the experimental design used to test consumer behaviors. In the meantime, testing hypotheses are linked to the measures in the experiment. Section 3.3 presents results of the study. Causal inference, robustness and manipulation

Table 8: Study 2_Experimental Procedure

Control	Treatment
Instruction and Hypothetical Practice Bid	Instruction and Hypothetical Practice Bid
	Participant Chooses 1 Food Item to Bid
Non-Hypothetical Bid for **3** Food Items - Round 1	Non-Hypothetical Bid for **1** Food Items - Round 1
Relaease **Qualitative** Information	Relaease **Qualitative** Information
Non-Hypothetical Bid for **3** Food Items - Round 2	Non-Hypothetical Bid for **1** Food Items - Round 2
Relaease **Quantitative** Information	Relaease **Quantitative** Information
Non-Hypothetical Bid for **3** Food Items - Round 3	Non-Hypothetical Bid for **1** Food Items - Round 3
Random Draw to Implement 1 bid out of **9**	Random Draw to Implement 1 bid out of **3**
Survey ans Settlement	Survey ans Settlement

check are also performed in this section. Section 3.4 provides discussion of the paper.

3.2 Experimental Design

3.2.1 Treatments and Experimental Procedure

This study used a between-subject experiment design. Grocery shoppers were recruited through existing mailing list of staff members at Cornell University and were guaranteed with $10 for participation. Experiment was conducted on an individual basis. Upon arrival, subjects were randomly assigned to 1 of 2 groups. In the control group, the subjects were invited to bid for 3 food items with different flavors but otherwise identical for 3 rounds consecutively, and were told 1 out of the 9 bids would be randomly chosen to be implemented at the end. In the treatment group, the subjects were invited to freely choose 1 flavor out of the 3 otherwise identical food items to bid for 3 consecutive rounds and were told 1 out of the 3 bids would be randomly chosen to be binding at the end. Table 8 provided a brief list of the experimental procedure for both control and treatment groups. A detailed procedure for each group was described as follows.

For the control group, as subjects walked into the room, they were told the experimenter was interested in consumer's food preference and evaluation. Subjects were each guaranteed $10 for completing the study. Subjects were also told the study involved non-hypothetical auctions for food items and based upon their bids and 2 lottery results, they had a chance to win one of the food items and pay the corresponding price. The procedure of the auction was then explained to subjects and a hypothetical practice round using a small piece of stationary (a small pack of BIC sticky notes) was conducted to make sure subjects understand the auction procedure. Subjects were also instructed

that they could ask the experimenter if there were any questions or concerns, but they were not allowed to talk to each other or share results or bids.

After going through a hypothetical practice auction using a non-food item and making sure subjects were clear about how the auction worked, the experimenter brought subjects to the real auction. 3 food items were presented and subjects were told that the whole auction involved making bids for the 3 food items in 3 rounds. In order to encourage subjects to be serious in making all 9 bids, a friendly reminder was given before they made any bid. The reminder said the auction for each item in each round worked in the same way as in the practice session and all of the 9 auctions were equally likely to be chosen and implemented. It was to the participant's benefit to treat each auction seriously as if it was the one that would determine final payoffs.

In each bidding round, subjects were asked 3 pairs of questions, one pair for each food item. The first question asked about food safety risk perception for the particular food item ("What do you think is the risk that people getting sick from eating XXX – the food item, where 1 means no risk at all and 10 means absolute risk?"). The second question is willingness to pay for the food item. Subjects were instructed to choose "Yes" (willing to) or "No" (not willing to) to purchase the food item at each of 10 listed prices, ranging from \$0.10 to \$1.00 with increment of \$0.10. Within each round, the order of bidding items were randomized in order to control for order effects on bidding behavior. Round 2 and 3 were conducted the same way as in round 1, except that before these two rounds, qualitative and quantitative information sessions regarding food safety risk and some food-borne pathogen were given to subjects. The content and details of the information were discussed in the next subsection.

When subjects finished their bids for all 3 rounds, the experimenter offered two lotteries to each subject. The first lottery (chips numbered from 1 to 9) determined which of the 9 bids was implemented. The second lottery (chips numbered from 1 to 10) determined the final price of the bidding item. Based upon bids and lottery results, subjects were then told individually whether they won the auction and if yes, what their final payoffs should be. Subjects were asked to finish a survey before they could get monetary payoffs. The survey included questions such as preference of the food items, eating habits, current sense of hunger, previous knowledge and awareness of food safety issues, risk perceptions and demographics such as age, gender, income and household composition, etc. Upon completion, subjects got their monetary settlements and the experiment

ended.

For treatment group, the procedure was the same as the control group except the auction rounds. In the treatment group, instead of bidding for all 3 listed food items, subjects were told to freely choose 1 food item out of the 3 to bid for 3 rounds. But same as in the control group, whenever the 3 listed food items came together in the instruction or choice options, the order of the 3 was randomized across subjects so as to control for order effects. A reminder was given to subjects that once they made their choice, they would bid only for the item they chose and would only have a chance to win the chosen item. In this case, in each bidding round, subjects were only asked 1 pair of questions regarding the chosen item. First was the risk perception and second was the willingness to pay. Same pieces of information (i.e. qualitative and quantitative) were provided before round 2 and round 3. Lottery choices were still used to determine final payoffs. The only difference was in the first lottery; chips numbered from 1 to 9 indicated 3 rounds rather than 9 separate bids, with 1 to 3 representing round 1, 4 to 6 representing round 2 and 7 to 9 representing round 3. Survey was also given to subjects before they could get monetary payoffs and ended the session.

3.2.2 Food Items, Information and Dissonance Inducing

The essential part in this study was to induce dissonance feelings and investigate consumers' behavioral responses. In addition to free choice (Treatment) and random assignment (Control) of bidding items, other critical conditions to induce dissonance are similarity among choice options and availability of conflicting information.

Chocolate candy bars, which were different only in flavors but otherwise identical, were used as bidding items in the study. The three different flavors were plain/original, peanuts and almonds. There were three major reasons or concerns in choosing bidding food items for the study. First, the items needed to be highly identical to one another with only one trivially different attribute (Different flavors in this study). In this sense, it was easier to make external intervention to the trivial attribution so as to induce dissonance feelings, while at the same time excluding all other potential compound impacts. Subjects in the treatment group picked their bidding items by preference or by some temporary and unconscious decisions. Exogenously offering food safety information regarding some of the flavors/food items afterwards induced dissonance feelings among those who chose the corresponding items. More discussion about the information is provided in the next few paragraphs

in this section.

Second, the selected food items were common and tempting in general. On the one hand, using food items that were commonly seen in the real life yielded results that could be easily generalized to most food purchase and consumption situations. One caveat was to choose common food items but less common brand names, since popular brands with high reputation were normally believed to be free of food safety issues and hence, could potentially impede the induction of dissonance feelings. On the other hand, choosing food items that were tempting enough attracted subjects and made them to be willing to bid some money on the food in the experiment, especially when it was the snack time in the middle or late afternoon. The changes in bidding behaviors and reactions to information driven by temptation and dissonance feelings could also provide implication for public health and food safety regulation.

Third, big gap between wholesale price and retail price was another merit. Large variation in price made subjects feel free to bid within the range so as to capture changings in risk perception, preference and feelings in the study. At the same time, since the large price range also exist in the real life, the subjects wouldn't feel it was too unrealistic to bid so differently across products and section rounds.

In addition to the selection of food items, providing conflicting information was another crucial part of the design. In this study, aflatoxin food-borne pathogen was selected. First of all, it was not so widely familiar to the public as other pathogens such as semolina or E. coli, etc. Subjects would be tested for their reactions in according to their self-reported individual awareness or knowledge of the pathogen. Secondly, food items were involved with different levels of aflatoxin risk. Among the selected food items in the study, peanut flavor candy bars were the ones with highest risk of aflatoxin, and almond flavor had the second highest risk, while plain/original flavor was generally believed to be free from the risk. Special attention would be given to the interaction between the choice of food items and the nature of risk involved. Lastly, aflatoxin was highly associated with corn and nut products, which took a substantial proportion of common foods. Testing public responses to pathogen of this type could generate rich policy implications.

During the experiment, food safety information was provided to subjects twice, once before bidding in the second round and once before bidding in the third round. The first information sheet included some general qualitative introduction about aflatoxin, how it was related to common food

items and how it was related to human health and potential sickness, etc. In addition, peanut and almond were pointed out to be highly relevant to this food-borne pathogen. In the second information sheet, quantitative information about aflatoxin concentrations detected in different types of products was provided. Comparisons were also made between peanut products and almond products. It was also made clear that whenever a product was detected with aflatoxin, concentrations in peanut products were roughly 1000 times more than that in almond products. These two information sessions were provided so as to induce dissonance feelings for those subjects who chose peanut or almond candy bars in the treatment group. Information was designed to be general and vague in the first session and was designed to be more clear and standardized in the second session . Impact of dissonance feelings, effectiveness of information and potential spillover effects can then be tested for both immediately related food items (peanut and almond candy bars) and indirectly involved item (plain flavor bars).

3.2.3 Auction Mechanism

In order to hold participants' decisions accountable, incentive-compatible auction mechanism was used in the study. In general, experimental auction has the advantage of putting participants in a situation with real goods and real money. Moreover, consumption requirement in the experiment would force individuals to put cognitive efforts into their bidding decisions (Fox et al, 1995). In particular, the Becker-Degroot-Marschak (BDM) mechanism was adopted here.

The BDM is a popular value elicitation mechanism, but it is not an auction per se as subjects do not bid against one another in a market environment. However, the fundamental logic is same as other auction mechanisms. With the BDM, a person submits a bid and purchases the good if his bid is greater than some randomly drawn price.

Other auction mechanisms that are commonly used in practice are English auction, 2nd price auction and (random) nth price auction. In brief, English auction is the most familiar mechanism to the general population, where people offer ascending bids until only one participant remains. The person wins the auction and purchases the good at the last offered bid amount. In 2nd price auction, subjects submits sealed bids, the highest bidder wins the auction and pays the 2nd highest bid. The nth price auction is a generalization of 2nd price auction, where people submit sealed bids and the (n-1) highest bidders win the auction and pay the nth highest bid amount. Random

nth price auction differs in the sense that the cutting rank n is randomly drawn, rather than being predetermined.

In principle, these auctions mechanisms yield bids equal to the true values and thus are equally efficient in eliciting preference and valuation. In practice, however, certain mechanism will have some relative advantages in certain context. The choice of mechanism for a specific study depends upon the particular purpose and the context.

In the induced value experiment setting, on average, the 2nd price and the nth price auction are more accurate than the BDM. In particular, 2nd price auction works better on-margin (i.e. for people with valuation near the market price), and the random nth price auction works better off-margin (i.e. for people with valuations far from the market price). The 2nd and nth price auctions have advantages when estimating the demand curve with market interactions among bidders (Shogren et al, 2001; Noussair et al; 2004 and Lusk and Rousu, 2006).

In the homegrown value experiment setting, however, all auction mechanisms tend to yield similar mean bids in the initial rounds, with 2nd price auction overbidding and nth price auction underbidding in the later rounds (Lusk et al, 2004b), suggesting an order effect and affiliation (peer effect). The BDM and English auction in general generate similar bids in all rounds (Rutstrom, 1998).

Further, BDM is the only mechanism that could be performed on an individual basis that does not require a group of subjects. It has been proved to be useful for eliciting values in field setting such as grocery stores (Lusk et al, 2001; Rousu et al, 2005). In this sense, BDM fits the purpose of this study best. A few considerations that need to be taken into account when conducting BDM are: First, lack of active market environment might be crucial in inducing economic rationality at the margin or at the individual level if arbitrage has enough power. Second, the WTP/WTA disparity that disappears in active market environment will still exist in the BDM. Third, the distribution for the random price generator needs to be chosen very carefully. It should be large enough (so that the bids will not be censored) but with limited range (so that reasonable bids have enough chance of winning), and still easier to be conveyed to the participants.

Specific to the BDM auction in this study, the procedure had been designed carefully to avoid the potential problems above. First, the nature of this study was not elicit the true market price for certain food items, but was to investigate the changings in valuations across rounds due to

some psychological biases. Hence, active market environment and on/off margin consumers were less important concerns. Further, the study was designed as one-side transaction, which was to let subjects make bids and purchase the food item when they won. So, there should not be any chance for arbitrage. Second, as mentioned earlier, since the interest of this study was the changes in the valuation across rounds and by the design, only the WTP was used, hence, the WTP/WTA disparity was irrelevant here. Third, careful instructions were made to the subjects before the auction and a hypothetical practice round was performed before the real auction for food items so as to make sure subjects understand the whole auction mechanism. The price range and random price generator device were also well controlled in the study to ensure 50% chances to win by expectation. More discussion was included in the result section.

To be more detailed, during the auction procedure, as subjects arrived for the study, they were told the purpose of this study was to investigate consumers' valuation of certain foods via auction mechanisms. The following auction procedure was then explained to them.

"Here is how the auction will proceed:

1. *First, you will indicate your willingness to pay for the auction item by circling "Yes" or "No" for EACH price listed on the bid sheet. NOTE: If you choose "Yes" at a certain price, your choices for all other prices lower than this price should also be "Yes". In contrast, if you choose "No" at a certain price, your choices for all other prices higher than this price should also be "No". The ONLY price that makes you switch from "Yes" to "No" is the MOST you are willing to pay for the item.*

2. *Then, the experimenter will randomly draw a number, 1 through 10; each indicates a corresponding price on your bid sheet. For example, if we draw the number 1, the price indicated is the 1st price on the list. In this case, it is $0.10. Importantly, all prices are equally likely to be drawn.*

3. *If your choice at the randomly drawn price is "Yes", then you win the auction, purchase the item at the price equal to the randomly drawn price.*

4. *If your choice at the randomly drawn price is "No", then you do NOT win the item."*

The subjects were instructed first to indicate their willingness to pay for the bidding food item by

choosing "Yes" or "No" for each of 10 prices listed on the bidding sheet. The price ranged from \$0.1 to \$1, with increment of \$0.1. A special explanation was given to subjects that there should be only one switching price from "Yes" to "No", which should be their highest willingness to pay for the food item. A lottery choice out of 10 chips each indicating one of the 10 prices was then offered to subjects. The instruction told them that if at their random drawn prices, they said "Yes" on their bidding sheets, they won the item and would actually pay the randomly drawn price. If instead, they said "No" at the randomly drawn price on their sheets, then they did not win. In order to help subjects further understand the procedure and make their bids accountable, some extra notes were given them explaining why it was to their benefits to bid exactly what the item was worth to them.

Important Notes:

- *In this auction, the Best Strategy is to bid EXACTLY what the item is worth to you. Consider the following: if you bid MORE than the item is worth to you, you may end up having to buy an item for more than you really want to pay. Conversely, if you bid LESS than the item is worth to you, you may end up not winning the auction even though you could have bought an item at a price you were actually willing to pay.*
- *It is acceptable to choose "No" at all listed prices.*

3.2.4 Testing Hypotheses

With the design of the study, the following hypotheses were supposed to be tested:

Hypothesis 1 : ***(Cognitive Dissonance):*** *Pre-committing to a certain item leads to more favorable evaluations, i.e. estimating a lower risk and/or bidding for a higher price (WTP).*

As mentioned previously, in the treatment group, subjects were instructed to freely choose 1 out of 3 chocolate candy bars to bid and would have a chance to win only the chosen flavor; whereas in the control group, subjects did not make any choice, bidding for all 3 items and having a chance to win a randomly chosen flavor. Comparing the two groups, subjects in the treatment group actually pre-committed themselves to 1 food item. Putting such restrictions before bidding made subjects at least being no better off, since they could always choose to bid the same for the item without pre-committing to it, while still keep the chance of winning other items at desired prices. According to

cognitive dissonance theory, subjects would in this case rate more favorably for the chosen item and less favorably for the rejected ones. Equation 1 and 2 were constructed to test these experimental responses.

$$Bid_{ij} = \alpha_0 + \alpha_1 Treatment_i + \alpha_2 Prefer_i + \alpha_3 Risk_{ij} + X_i\beta + \sum_{k=2}^{3} \mu_k + \epsilon_{ij} \tag{1}$$

$$Risk_{ij} = \delta_0 + \delta_1 Treatment_i + \delta_2 Prefer_i + X_i\beta + \sum_{k=2}^{3} \mu_k + \epsilon_{ij} \tag{2}$$

In the model, Bid_{ij} is the bid made by subject i in informational round j, where $j = 1$, 2 and 3. Similarly, $Risk_{ij}$ is the risk perception reported by subject i in informational round j. $Treatment_i$ and $Prefer_i$ are dummy variables, where $Treatment_i$ equals 1 for treatment, and 0 otherwise; $Prefer_i$ equals 1 if the subject indicated he/she preferred the corresponding food item, and 0 otherwise. μ_k represents the informational round fixed effect, round 2 or round 3, as oppose to the first no information round. X_i is a set of control variables. Note that Equation 1 and 2 apply to all 3 flavors, with postfixes *_pl* (*Plain*), *_pe* (*Peanut*) and *_al* (*Almond*) respectively.

Testing of $H1$ focused on the estimated parameters of α_1 and δ_1 for each flavor. Positive estimates of α_1 and negative estimates of δ_1 would support $H1$. In addition, preference of the food item was controlled in the model to get an exclusive estimation of the treatment effect. In general, preference yielded higher bids, indicating consumers being willing to pay a positive premium for the food they like. Only after controlling for preference could one say the additional jump in bids between treatment and control group was due to psychological temptations. The effect of preference on risk perception could be more controversial. In normal cases, the effect should be non-positive. Zero (or insignificant) for rational agents, since they could objectively perceive the risk regardless of their preference. For nonrational agents instead, the effect could be negative, indicating a self-justification/self-compliance tendency to reduce dissonance feelings. (However, for relatively more exotic and less familiar food, the relationship was undetermined, since consumers might prefer the food just because they could enjoy the fun of risk. An endogenous problem would bias the result. Fortunately, this concern was less relevant to this study, since the tested food items were commonly seen almost every day and everywhere.)

Hypothesis 2 : ***(Risk Aversion):*** *When being exposed to relevant risky information about common item, people in general increase the risk perception and decrease the WTP.*

Following from the above discussion about risk perception and preference, $H2$ was to test that, in general, consumers perceived risk as a bad attribute and tried to stay away from it when dealing with common items. In the study, participants were instructed to bid for candy bars (common food items) in three rounds. Before round 2 and round 3, information regarding potential food-borne risk involved with food items was revealed to them. According to Equation 1 and 2, $H2$ would be supported if the fixed effects μ_k, $k = 1, 2$, were negative in Equation 1 and positive in Equation 2. In addition, in Equation 1 the effect of risk perception on bids, that is α_3, being negative would further support the hypothesis.

$H2$ could also be tested by checking the average changes of bid and risk perception across rounds, shown in Equation 3 and 4.

$$dBid_i^{jk} = \gamma_0 + \gamma_1 Treatment_i + \gamma_2 Prefer_i + \gamma_3 Risk_i^j + \gamma_4 dRisk_i^{jk} + X_i\beta + \epsilon_i \tag{3}$$

$$dRisk_i^{jk} = \theta_0 + \theta_1 Treatment_i + \theta_2 Prefer_i + \theta_3 Risk_i^j + X_i\beta + \epsilon_i \tag{4}$$

In Equation 3 and 4, $Treatment_i$, $Prefer_i$ and X_i are the same as in Equation 1 and 2. $Risk_{ij}$ is the risk perception of individual i in round j, where $j = 1, 2$ and 3. $dBid_i^{jk}$ is the change of bid for individual i from round k to round j, where $j, k = 1, 2$ and 3, and $k < j$. Similarly, $dRisk_i^{jk}$ is the change of risk perception for individual i from round k to round j, where $j, k = 1, 2$ and 3, and $k < j$. Same as the previous two equations , Equation 3 and 4 apply to all 3 flavors , with postfixes *_pl* (*Plain*), *_pe* (*Peanut*) and *_al* (*Almond*) respectively.

Testing of $H2$ is now equivalent to testing the intercept terms γ_0 being significantly negative and θ_0 being significantly positive. Further, the model controls preference and the absolute level of risk perception. Detailed discussion will be given in the result session.

Hypothesis 3 : ***(Confirmatory Bias):*** *The tendency to reduce dissonance feelings will (i) mitigate (or attenuate) the impact of negative information and (ii) amplify (or exaggerate) the impact of positive information.*

When people experience cognitive dissonance, a natural tendency to reduce dissonance feeling is to selectively pay attention to (or even make up) information that is in line with previous behaviors and to intentionally overlook the information that generates the confliction (Frey, 1986). In the study, food-borne risk information about Aflatoxin was provided to participants before the second and third rounds of auction. Participants' reactions in the control group served as a benchmark. According to $H2$, risky information would increase risk perceptions of involved food items, i.e. peanut and almond flavored candy bars, while leaving the risk perception of uninvolved item, i.e. plain flavored candy bar unchanged.

However, based on choices in the treatment group, the participants reacted differently to support $H3$. First, if the subjects who pre-selected peanut or almond flavor were less willing to increase their risk perceptions, then $H3-(i)$ was supported. In Equation4 one would expect negative θ_1 for both peanut ($\_pl$) and almond ($\_al$) equations between information round (round 2 or 3) and initial round (round 1). Further, information before the third round auction quantitatively compared Aflatoxin food-borne risk between peanut and almond. The fact that although the food-borne risk was typically high in almond products, the concentrations of Aflatoxin detected from them was only about 1/1000 of that from peanuts was mentioned in the information sheet. While subjects in the control group would still pick up the information that an almond product was typically high in Aflatoxin risk and would keep the risk perception roughly unchanged, those who pre-committed to the almond candy bar in the treatment group would focus more on the fact that almond was less risky than peanut and hence slightly decrease their risk perceptions from the 2nd round to the 3rd round. If this was the case, $H3-(ii)$ would be supported. Checking Equation 4 for the almond equation with $j=3$ and $k=2$, estimate of θ_1 should be slightly positive.

Hypothesis 4 : ***(Sticky Behavior):*** *The taste of consistency makes people less likely to change their behaviors, even though it is to their benefits to do so.*

Even though subjects increased their risk perceptions, they would be more reluctant to change their bids in the treatment group, in which case they pre-committed themselves to one certain item. This taste of consistency would be supported by positive estimates of γ_1 in Equation 3. In real life, this sticky behavior caused by cognitive dissonance and confirmation bias made consumers less responsive to food-safety information and keep consuming certain products even when facing high

risk.

3.3 Results

3.3.1 Summary Statistics

Ten sessions of experiments were conducted in mid-late afternoons within two days. A total of 116 subjects participated in the study, 42 of whom were randomly selected to the control group and 74 to the treatment group. Among those in the treatment group, 18 subjects freely chose to bid for a plain flavor candy bar, 20 chose peanut flavor and the remaining 36 chose almond flavor. Three general groups of responses were collected for all the participants: 1) risk perceptions and biddings in the auction games; 2) objective demographics and household background; and 3) subjective beliefs, including past eating habits, knowledge, awareness and judgment of food-safety issues, etc.

Table 9 listed the demographics and household background for both treatment and control groups. In general, the two groups were well balanced for the characters such as basic body measures, social demographics, household composition and dietary habits, etc. The age of participants in both groups ranged between 18 and 52, with an average of about 41 years old. The mean level of height and BMI were about 66 inches and 23 respectively. 45.69% were White/Caucasian. The majority of the subject pool had at least 2-year college education and was currently employed either part-time or full-time, with an average annual household income of \$50K to \$75K. On average, the participants came from families of 3-4 people with 1 kid under 18 years old. Roughly 70% of the participants claimed themselves as primary grocery shoppers for their families. In terms of eating habit, 7% said they were currently on diet and 29% had a past experience to be on diet. At the time of study, it was about 3.27 hours before the participants ate any food.

The main responses in the study were risk perceptions and bids reported by participants during the auction. Table 10 presented summary statistics with F-test between treatment and control groups.

The top panel reports results of risk perceptions for 3 flavors in 3 rounds. First of all, control and treatment groups were not different from each other in estimating risk in the initial round for the 3 flavors. This was the risk perception when no information regarding food-borne illness was shown to participants. On average, they estimated 2 out of 10-point-scale for risk with plain

Table 9: Study 2_Summary Statistics of Demographics

	Control					Treatment				
Variable	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Age	40	41.13	2.57	18	52	73	42.16	6.08	18	50
Gender	40	0.48	0.51	0	1	73	0.33	0.47	0	1
Height	40	66.93	4.83	57	79	73	65.88	3.59	56	76
BMI	39	22.75	3.21	17.92	32.61	73	22.95	3.62	17.64	34.54
HrsBfrEat	40	2.84	2.28	0.2	14	73	3.51	2.86	0	17
Diet	40	0.43	0.68	0	2	73	0.45	0.60	0	2
Marriage	40	1.13	0.46	1	3	73	1.08	0.36	1	3
Race	40	2.05	0.96	1	4	73	2.00	1.20	1	6
Edu	40	3.23	1.03	1	6	73	3.47	1.03	2	6
Employ	40	4.95	1.71	1	6	73	4.41	2.07	1	6
Income	40	3.98	2.41	1	7	73	4.21	2.38	1	7
Families	40	3.65	2.18	0	11	73	3.04	1.15	1	6
Children	40	0.55	0.90	0	4	73	0.42	0.66	0	3
Shopper	40	0.68	0.38	0	1	73	0.70	0.46	0	1

Table 10: Study 2_Summary Statistics of Risk Perception and WTP by Round

	Control					Treatment						
Variable	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max	F-test	P(>F)
Risk_plain1	42	2.38	1.71	1	8	18	2.22	1.44	1	7	0.119	0.732
Risk_plain2	42	2.12	1.40	1	6	18	2.06	0.87	1	4	0.032	0.860
Risk_plain3	42	2.12	1.35	1	6	18	2.00	0.84	1	4	0.120	0.730
Risk_peanut1	42	3.88	2.10	1	10	20	5.10	2.65	2	10	2.427	0.126
Risk_peanut2	42	5.76	2.28	2	10	20	5.90	1.97	3	9	0.031	0.862
Risk_peanut3	42	6.33	2.46	2	10	20	7.00	1.78	4	9	0.647	0.425
Risk_almond1	42	3.10	1.53	1	7	36	3.03	2.34	1	10	0.023	0.879
Risk_almond2	42	5.14	1.97	2	9	35	4.17	2.53	1	10	3.592	0.062
Risk_almond3	42	5.17	2.08	2	10	35	3.96	2.58	1	10	5.172	0.026
Bid_plain1	42	0.52	0.31	0	1	18	0.68	0.24	0.1	1	3.826	0.055
Bid_plain2	42	0.54	0.30	0	1	18	0.67	0.24	0.1	1	2.747	0.103
Bid_plain3	42	0.53	0.30	0	1	18	0.66	0.24	0.1	1	2.481	0.121
Bid_peanut1	42	0.54	0.27	0	1	20	0.78	0.14	0.6	1	7.012	0.011
Bid_peanut2	42	0.39	0.27	0	0.9	20	0.74	0.13	0.5	1	15.490	0.000
Bid_peanut3	42	0.32	0.27	0	0.9	20	0.69	0.20	0.2	1	16.620	0.000
Bid_almond1	42	0.62	0.27	0	1	36	0.65	0.27	0.1	1	0.297	0.587
Bid_almond2	42	0.45	0.27	0	0.9	36	0.58	0.28	0.1	1	4.473	0.038
Bid_almond3	42	0.42	0.28	0	1	36	0.56	0.28	0.1	1	4.596	0.035

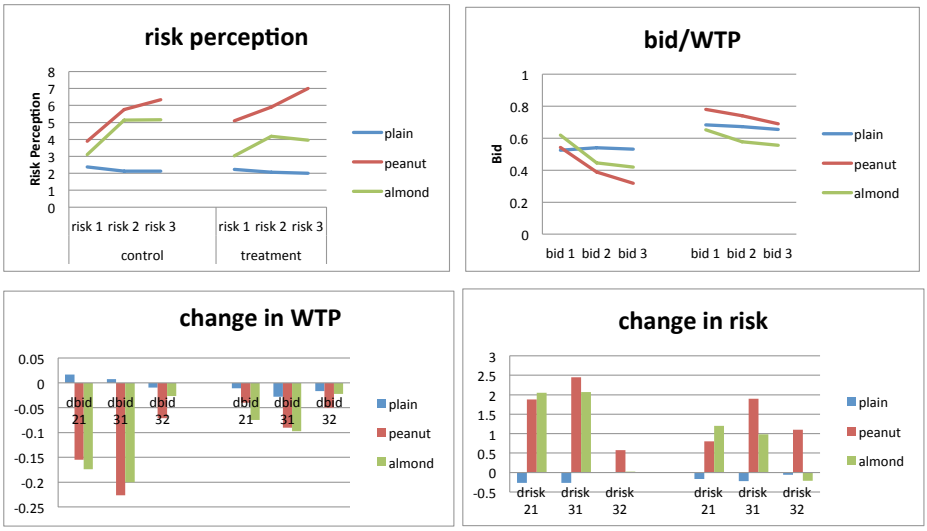
flavor, 3/10 with almonds, and 4.5/10 with peanuts. The treatment began to play a role in affecting risk perception when food safety information was provided in the 2nd and 3rd round. Since the Aflatoxin food-borne risk was highly related to peanut and almond products, but was almost not relevant to plain flavor candy bar, one would expect risk perception for peanut and almond flavored candy bars increased and risk perception for plain flavor remained unchanged. While this was the case, changings in risk perception were not the same between control and treatment groups. For almond flavor, control groups estimated much higher risk after reading information than treatment groups (5.14/10 for control vs. 4.17/10 for treatment, F=3.59, P>F=0.06 in round 2; and 5.17/10 for control vs. 3.96/10 for treatment, F=5.17, P>F=0.03 in round 3). For peanut flavor, though participants in the treatment estimated a slightly higher risk in the first round, their changes in the 2nd and 3rd rounds were much lower than those of control group. For example, on average, the increase in risk perception was 1.9/10 in control and only 0.8 in treatment (F=3.06, P>F=0.00) from round 1 to round 2. These differences between treatment and control group were first signals supporting cognitive dissonance and confirmation bias. More discussion is provided in later part of this section.

The bottom panel of Table 10 listed descriptive information about biddings. Based on the auction mechanism, the bids revealed participants' willingness to pay (WTP) for each food item in each informational round. One would easily see from Table 3 that participants in treatment bid higher prices for their pre-selected food items almost in all rounds, i.e. for plain flavor in round 1, 52 cents in control vs. 68 cents in treatment, F=3.83, P>F=0.06; for peanut flavor in round 2, 39 cents in control vs. 79 cents in treatment, P>F=0.00. Further, bids in treatment group roughly remained the same across all 3 rounds, while in the control group bids for peanut and almond flavor bars were clearly decreased in the 2nd and 3rd round.

Figure 1 provided a straightforward description of these behavioral responses.

In addition to the above two categories, a third category of responses – beliefs and judgments was also collected in the post-study survey. More investigation for these responses and their interactions with the risk perceptions and bidding behaviors is provided later.

Figure 1: Study 2_Risk Perception and Bid by Information Round & Treatment



3.3.2 Cognitive Dissonance

Cognitive dissonance hypothesis ($H1$) was supported if participants in the treatment group made higher bids in absolute value than those in the control group. The top panel of Table 11 listed the results of regressing 9 bids (3 flavors in 3 rounds) on the treatment dummy. Initially, participants in the treatment group bid higher prices for all the 3 flavors. 16 cents higher for plain (t=1.96, P=0.06), 24 cents higher for peanut (t=3.96, P=0.00) and about 3 cents higher for almond, though not statistically significant. In the 2nd and 3rd round, risk information was provided to participants. If the risk information involving peanut and almond can further intensify the dissonance feelings, one would expect the treatment effect on bid to be enlarged for peanut and almond flavors in these two rounds. The results were in line with this hypothesis. Peanut flavor candy bars were bid 35 cents and 37 cents higher in the 2nd round and the 3rd respectively. Almond flavor was bid 13 cents and 14 cents higher and the significant level was 5%. For plain flavor, since it was not involved in the risk information, the difference in bid between control and treatment disappeared in the last two rounds, which also supported the dissonance hypothesis on the flip side. That was, initial dissonance due to free choice (/pre-commitment) could be reduced by any (ambiguous) information that seemed to support the earlier choice behavior.

To further verify treatment effects, regressions for each flavor were run using only those who claimed they loved the corresponding flavor. For each flavor, this subset of the sample pool included

Table 11: Study 2_Average Treatment Effects on Bid - All & Preferred Only

All	1	2	3	4	5	6	7	8	9
VARIABLES	Bid_pl1	Bid_pl2	Bid_pl3	Bid_pe1	Bid_pe2	Bid_pet3	Bid_al1	Bid_al2	Bid_al3
Treatment	0.160*	0.132	0.125	0.237***	0.352***	0.373***	0.034	0.133**	0.137**
	(1.956)	(1.657)	(1.575)	(3.659)	(5.448)	(5.477)	(0.545)	(2.115)	(2.144)
Constant	0.524***	0.540***	0.531***	0.543***	0.388***	0.317***	0.619***	0.445***	0.419***
	(11.730)	(12.410)	(12.250)	(14.750)	(10.580)	(8.180)	(14.730)	(10.460)	(9.687)
Obs	60	60	60	62	62	62	78	78	78
preferred only	1	2	3	4	5	6	7	8	9
VARIABLES	Bid_pl1	Bid_pl2	Bid_pl3	Bid_pe1	Bid_pe2	Bid_pet3	Bid_al1	Bid_al2	Bid_al3
Treatment	0.057	0.006	0.036	0.347***	0.407***	0.457***	-0.097	0.015	0.031
	(0.630)	(0.069)	(0.419)	(4.699)	(6.257)	(5.591)	(-1.310)	(0.189)	(0.368)
Constant	0.627***	0.667***	0.620***	0.433***	0.333***	0.233***	0.750***	0.562***	0.525***
	(9.428)	(11.260)	(9.887)	(7.073)	(6.175)	(3.440)	(12.150)	(8.362)	(7.599)
Obs	33	33	33	29	29	29	52	52	52

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

all participants in the treatment, and the part of participants in the control who reported in the survey that they prefer the certain flavor to the other two alternatives. Based on self-reported preferences, final subsets had 33 observations for plain flavor, 29 for peanut and 52 for almond. This partition of data could help to exclude the effect of preference on bidding premiums between two groups, and get a much cleaner estimation of the dissonance effect.

The bottom panel of Table 11 presented the results using these sub-samples. The bidding premium was decreased to only 6 cents for plain flavor in the initial round, and 2 to 3cents for almond flavor in the two informational rounds. This implied preference did perform a role in making participants bid higher. However, after excluding preference effects, bidding premiums for peanut candy bars remained to be 30-40 cents with 1% significant level. Existence of cognitive dissonance was established even after controlling for preference effects.

Table 12 listed the estimated average treatment effects (ATE) on bidding premiums with different combinations of the control variables. As shown in the table, treatment effects were stable while controlling for risk perceptions and preferences. Further, effects of preference on bidding for all 9 scenarios within control group were listed on the bottom. Subtracting these numbers from the treatment effects using all observations (top panel of Table 12) could also yield exclusive estimations of dissonance effects on bidding behaviors (so long as treatment was assigned randomly, which was guaranteed by the design of the study).

Table 12: Study 2_ATE on Bid - Robustness Check

VARIABLES	1 Bid_pl1	2 Bid_pl2	3 Bid_pl3	4 Bid_pe1	5 Bid_pe2	6 Bid_pet3	7 Bid_al1	8 Bid_al2	9 Bid_al3
ATE no control	0.160*	0.132	0.125	0.237**	0.352***	0.373***	0.034	0.133**	0.137**
	(1.956)	(1.657)	(1.575)	(2.648)	(3.936)	(4.076)	(0.545)	(2.115)	(2.144)
ATE ctrl for risk	0.159*	0.131	0.123	0.227***	0.356***	0.403***	0.034	0.131**	0.117*
	(1.929)	(1.634)	(1.546)	(3.381)	(5.673)	(6.359)	(0.551)	(2.061)	(1.823)
ATE ctrl for risk & pref.	0.052	0.005	0.03	0.307***	0.405***	0.506***	-0.054	0.066	0.064
	(0.551)	(0.061)	(0.326)	(3.670)	(5.026)	(6.218)	(-0.787)	(0.930)	(0.880)
Obs.	60	60	60	62	62	62	78	78	78
ATE w/ preferred only	0.057	0.006	0.036	0.347***	0.407***	0.457***	-0.097	0.015	0.031
	(0.630)	(0.069)	(0.419)	(4.699)	(6.257)	(5.591)	(-1.310)	(0.189)	(0.368)
ATE ctrl for risk	0.053	0.003	0.033	0.329***	0.417***	0.491***	-0.097	0.029	0.022
	(0.576)	(0.041)	(0.375)	(4.218)	(6.340)	(5.830)	(-1.300)	(0.359)	(0.264)
Obs.	33	33	33	29	29	29	52	52	52
Diff due to preference	0.163*	0.199**	0.14	-0.16	-0.092	-0.125	0.212***	0.196**	0.179**
	(1.731)	(2.254)	(1.545)	(-1.624)	(-0.906)	(-1.228)	(2.714)	(2.539)	(2.160)
Obs.	40	40	40	40	40	40	40	40	40

3.3.3 Manipulation Check

Manipulation in the study was the information regarding Aflatoxin food-borne risk provided to participants before the 2nd and the 3rd round of auction. Peanut products were highly involved with this risk, followed by almond products. Those plain flavored chocolate candy bars were believed to be unrelated. Validity of this manipulation could be verified by checking reported risk perceptions across 3 rounds for all 3 flavors.

On average, participants reported 2.08 out of 10 points of risk associated with plain flavored candy bars, 5.3 out of 10 with peanut flavor and 4.67 out of 10 for almond flavor. Testing the equality between any pair of means was rejected, suggesting almond flavor was perceived higher in risk than plain flavor (t=3.58, P=0.00) and risk perception of peanut flavor was the highest among the three (peanut vs. almond, t=4.86, P=0.00). The manipulation of different risk perceptions across product was established successfully.

Further, Table 13 showed the average levels of participants' self-reported risk perceptions in 9 cases. As shown in the constant row, risk perceptions for peanut increased from 4.4 in the 1st round to 5.8 in the 2nd and 6.5 in the 3rd; risk perception for almond also increased from 3.3 in the initial round to about 5.2 in the later rounds, while for plain flavor, risk perception remained constantly at around 2. The changes in risk perceptions across rounds were also tested using Equation 4. The average change of risk perception corresponded to the intercept term θ_0 in Equation 4. As shown

Table 13: Study 2_Independence of Risk Perception on Treatment & Preference

	1	2	3	4	5	6	7	8	9
VARIABLES	Risk_pl1	Risk_pl2	Risk_pl3	Risk_pe1	Risk_pet2	Risk_pe3	Risk_al1	Risk_al2	Risk_al3
Treatment	-0.397	-0.122	-0.292	1.323	0.385	1.292	0.074	-0.831	-1.174*
	(-0.701)	(-0.290)	(-0.706)	(1.603)	(0.483)	(1.587)	(0.142)	(-1.370)	(-1.883)
Prefer	0.411	0.187	0.339	-0.293	-0.311	-0.877	-0.402	-0.252	-0.142
	(0.789)	(0.485)	(0.893)	(-0.375)	(-0.411)	(-1.138)	(-0.742)	(-0.403)	(-0.221)
Constant	2.246***	2.005***	1.973***	4.041***	5.795***	6.497***	3.311***	5.226***	5.257***
	(6.836)	(8.209)	(8.225)	(9.943)	(14.780)	(16.210)	(8.739)	(11.950)	(11.700)
Obs	57	57	57	60	60	60	76	75	75
R-squared	0.013	0.004	0.016	0.054	0.004	0.043	0.008	0.046	0.069

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 14: Study 2_ATE on Changes in Risk Perception - All & Preferred Only

All	1	2	3	4	5	6	7	8	9
VARIABLES	dRisk_pl21	dRisk_pl31	dRisk_pl32	dRisk_pe21	dRisk_pe31	dRisk_pe32	dRisk_al21	dRisk_al31	dRisk_al32
Treatment	0.095	0.04	-0.056	-1.081***	-0.552	0.529*	-0.848***	-1.086***	-0.238
	(0.283)	(0.127)	(-0.454)	(-3.058)	(-1.102)	(1.775)	(-2.696)	(-2.972)	(-1.010)
Constant	-0.262	-0.262	0	1.881***	2.452***	0.571***	2.048***	2.071***	0.024
	(-1.422)	(-1.532)	(-0)	(9.367)	(8.615)	(3.379)	(9.661)	(8.412)	(0.150)
Obs	60	60	60	62	62	62	77	77	77

Preferred only	1	2	3	4	5	6	7	8	9
VARIABLES	dRisk_pl21	dRisk_pl31	dRisk_pl32	dRisk_pe21	dRisk_pe31	dRisk_pe32	dRisk_al21	dRisk_al31	dRisk_al32
Treatment	0.3	0.111	-0.189	-0.978*	-0.322	0.656	-0.800*	-1.077**	-0.277
	(0.596)	(0.244)	(-1.164)	(-2.033)	(-0.392)	(1.177)	(-1.841)	(-2.087)	(-0.781)
Constant	-0.467	-0.333	0.133	1.778***	2.222***	0.444	2.000***	2.063***	0.063
	(-1.256)	(-0.993)	(1.113)	(4.452)	(3.254)	(0.961)	(5.556)	(4.826)	(0.213)
Obs	33	33	33	29	29	29	51	51	51

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

in Table 14, for example, the average change in risk perception for peanut candy bar was 1.88 (t= 9.37, P=0.00) in round 2 as opposed to round 1 and the change for almond candy was 2.07 (t=8.41, P=0.00) in round 3 compared to round 1. Hence, the manipulation of different risk perceptions across informational rounds was also proved to be effective. $H2$ is supported.

In addition, risk perceptions were shown to be unrelated with preferences (Table 13). This implied that participants were to some extent rational enough to objectively estimate risk and were not impacted by their own preferences. This finding validated Equation 1 through 4, which used risk perception and preference as two independent control variables to estimate treatment effects.

During the study, wherever a list of the 3 food items was mentioned to participants, the order of the 3 was randomized so as to make sure there was no confounded order effect. Further, regressions

were run to double check. For the control group, participants' bids were not significantly related to the order of the food being mentioned. For the treatment group, order of the food items did not predict participants' choices, either.

3.3.4 Confirmation Bias

Though participants did respond to the information provided to them, the magnitude of reaction was not the same between treatment and control groups. These findings brought us to the investigation of $H3$, confirmatory bias. Confirmatory bias was defined as a natural tendency to selectively pay attention to relevant information so as to reduce the dissonance feelings. To be more detailed, when participants tried to overlook the available information that was conflicting to their previous choices and/or became less sensitive to it, they were experiencing confirmatory bias. On the flip side, being sensitive to some irrelevant information or interpreting information as supporting evidence to the previous behaviors were also signals of confirmatory bias.

Changes in risk perceptions across rounds in Table 14 supported the above hypothesis ($H3$). Columns 4-6 reported regression results of changes in risk perception for peanut flavor between the 2nd and the 1st round (i.e. $dRisk\_pe^{21}$), the 3rd and the 1st round and the 2nd and the 3rd round respectively. Similarly, columns 7-9 were the results for almond flavor.

Take the peanut flavor candy bar as an example, the control group people on average increased their risk perception by 1.881 points after reading the first piece of information in the 2nd round auction (Column 4 in Table 8). However, those in the treatment only increased their perception by 0.8 points (=1.881-1.081), implying that they were more reluctant to respond to the risk information regarding the food item they just pre-committed to. Same thing happened to the almond flavor, too. The control group increased risk perception by 2.048 (Column 7), whereas the increase in the treatment group was only about 1.2 (=2.048-0.848).

Proportion of perception updating that happened in the first stage (i.e. from the 1st round to the 2nd round) also demonstrated difference in sensitivity to information between two groups. Take peanut flavor as an example, in the control group, risk perception increased 2.452 in total from the beginning to the end (i.e. difference between 1st and 3rd round), and increased 1.881 (points) immediately in the 2nd round, which was the first chance when the participants could update their perceptions. In this sense, participants made 76.71% (=1.881/2.452) of updating in the first stage.

In contrast, in the treatment group, the total increase in risk perception was about 1.9 (=2.452-0.552), and about 0.8 (=1.881-1.081) in the 2nd round. The immediate updating proportion was only 42.11% (=0.8/1.9). This result implied that participants in the treatment group were more reluctant to respond to information about negative attributes of the food they chose earlier.

For almond flavor candy bars, participants in the treatment group increased their risk perceptions by only 0.985 (=2.071-1.086) in total, compared with 2.071 in the control group. But interestingly, as early as in the 2nd round, the increase was 1.2 (=2.048-0.848) in the treatment group. This implied that in the 3rd round, participants even lowered risk perceptions by roughly 0.215 (=1.2-0.985) points in the treatment group. According to the information being provided in the 3rd round, while people in the control group still picked up the main idea that almond products were high in Aflatoxin food-borne risk, those in the treatment obviously were more attracted by the signal that almond products had a much lower detected concentration of Aflatoxin than peanut products and thus, interpreted it as a favorable evidence to modify their beliefs. The above findings supported H3, that people selectively pay attention to information so as to reduce dissonance feelings. The bottom part of Table 14 controlled for preference by using the subset of the participants who claimed they preferred the food for regressions, just as in Table 11 (Section 8.2). Results showed that after restricting observations to only those who had strong preference to the food in both groups, the major findings about confirmatory bias still remained.

3.3.5 Sticky Behavior

While participants updated their beliefs differently after reading pieces of information, it is interesting to see if their purchase tendency was impacted in the same way. By$H4$, consumers would have a taste of consistency and hence, be less willing to change their purchase behaviors. In the study, the participants in the treatment made their choices explicitly in the first place, but people in the control group did not. If we can find the participants in the treatment group were less willing to change their bids after reading risk information, then H4 was supported.

Table 15 showed the results of treatment effects on changes in bid. Similar as in Table 14, the top panel used all observations and the bottom panel used only those who claimed they preferred the food to control the effect of preference. As shown in the table, for peanut flavor, participants in the control group decreased their bids by 16 cents in the 2nd round and 7 cents in the 3rd

Table 15: Study 2_ATE on Changes in Bid - All & Preferred Only

All	1	2	3	4	5	6	7	8	9
VARIABLES	dBid_pl21	dBid_pl31	dBid_pl32	dBid_pe21	dBid_pe31	dBid_pe32	dBid_al21	dBid_al31	dBid_al32
Treatment	-0.028	-0.035	-0.007	0.115***	0.136***	0.021	0.099***	0.103**	0.004
	(-0.913)	(-0.853)	(-0.305)	(3.252)	(2.833)	(0.646)	(2.663)	(2.423)	(0.164)
Constant	0.017	0.007	-0.01	-0.155***	-0.226***	-0.071***	-0.174***	-0.200***	-0.026
	(1.000)	(0.319)	(-0.742)	(-7.723)	(-8.283)	(-3.789)	(-6.894)	(-6.942)	(-1.592)
Obs	60	60	60	62	62	62	78	78	78

Preferred only	1	2	3	4	5	6	7	8	9
VARIABLES	dBid_pl21	dBid_pl31	dBid_pl32	dBid_pe21	dBid_pe31	dBid_pe32	dBid_al21	dBid_al31	dBid_al32
Treatment	-0.051	-0.021	0.03	0.06	0.11	0.05	0.113**	0.128**	0.015
	(-1.022)	(-0.326)	(0.971)	(1.216)	(1.574)	(1.036)	(2.116)	(2.116)	(0.542)
Constant	0.04	-0.007	-0.047**	-0.100**	-0.200***	-0.100**	-0.188***	-0.225***	-0.037
	(1.083)	(-0.139)	(-2.044)	(-2.440)	(-3.447)	(-2.496)	(-4.239)	(-4.478)	(-1.600)
Obs	33	33	33	29	29	29	52	52	52

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

round. However, subjects in the treatment group decreased their bids by only 4 cents in the 2nd round and 5 cents in the 3rd round. The differences between treatment and control were significant at 1% level. In terms of proportion, the control group made 69.57% (=16 / (16+7)) of the total changing immediately after reading the first piece of information, whereas the treatment group made only 44.44% (=4 / (4+5)) of the changing in the first available chance. Taking absolute values of initial bids into account, the same information regarding food-borne illness of the same food item yielded about 42.36% (=0.23/0.543) decrease in willingness to pay (WTP) for those who did not pre-committed to the food, but only about 11.54% decrease in WTP for those who made their pre-commitment. While one could say loyal customers were less likely to be affected by negative news of their beloved products, they were also less sensible and less responsive in picking up crucial information when health or even safety issues were involved. Same patterns were also found for almond flavor in the table.

In addition, it was sensible to check the proportion or number of participants who changed their preferences after the study and try to link this to the changing in risk perceptions and bidding behaviors could be interesting.

Table 16 above listed absolute numbers and percentages of participants' claimed preferences. In the control group, roughly 9.52% (=26.19%-16.67%) of participants' switched away from peanut products after the study, and 7.15% (=40.48%-33.33%) for almond products. In comparison, the switching percentages in the treatment group were only 5.41% (=28.37%-22.97%) and 4.95%

Table 16: Study 2_Preference Changes Before & After Study by Flavor

		Prefer Plain	Prefer Peanut	Prefer Almond	Total
Control	Before #	14	11	17	42
	%	33.33	26.19	40.48	100
	After #	21	7	14	42
	%	50	16.67	33.33	100
Treatment	Before #	21	21	32	74
	%	28.38	28.38	43.24	100
	After #	28	17	29	74
	%	37.84	22.97	39.19	100

(=43.24%-39.19%) respectively. Evidence suggested that after pre-committing to some certain food (in the treatment group), dissonance feelings made subjects less willing to change their behaviors and/or more inclined to stick to their previous behaviors.

3.3.6 Habits, Risk Perceptions and Causal Effects (Identification and Validation)

After talking about how participants in the treatment group behaved differently from those in the control group in making and changing their bids (WTP: willingness to pay), this section starts to talk about what were the main reasons that caused these differences. In the study, there were two major reasons. One was long-term preference and the other was risk perception manipulated by the experimental design. The following two sub-sections discussed how these two factors interacted with psychological biases and finally impacted behaviors.

Determinants of Absolute WTP Table 17 listed partitioned regressions of bid for all 9 scenarios (3 flavors in 3 rounds). Risk perceptions and preference were controlled in these regressions.

First of all, preference did have significant impact on participants' bids. For plain and almond flavor, preference yielded about 15 cents more in WTP. This preference premium reduced the treatment effects to about 5 cents for plain flavor in the initial round and 6 cents for almond flavor in the two informational rounds. Previously, as shown in Table 11, these treatment effects were 16 cents for the former and 13 cents for the latter. However, one could not merely say these preference interpretations disprove the dissonance argument that had been made earlier. Observing preference premiums across (information) rounds within each flavor, one could find for plain flavor, the premium increased from the first round to the second round and for almond flavor, the premium decreased. One interpretation of these findings was that the manipulation of risk perception trig-

Table 17: Study 2_ATE on Bid - Causal Inference

	1	2	3	4	5	6	7	8	9
VARIABLES	Bid_pl1	Bid_pl2	Bid_pl3	Bid_pe1	Bid_pe2	Bid_pet3	Bid_al1	Bid_al2	Bid_al3
Treatment	0.052	0.005	0.03	0.307***	0.405***	0.506***	-0.054	0.066	0.064
	(0.551)	(0.061)	(0.326)	(3.670)	(5.026)	(6.218)	(-0.787)	(0.930)	(0.880)
Risk	-0.002	0.001	0.003	0.003	-0.028**	-0.048***	0.011	-0.011	-0.021
	(-0.0922)	(0.025)	(0.089)	(0.263)	(-2.128)	(-3.712)	(0.696)	(-0.795)	(-1.577)
Prefer	0.155*	0.186**	0.13	-0.131*	-0.096	-0.166**	0.175**	0.145**	0.126*
	(1.799)	(2.268)	(1.551)	(-1.686)	(-1.259)	(-2.181)	(2.454)	(2.021)	(1.726)
Constant	0.472***	0.471***	0.478***	0.573***	0.589***	0.670***	0.519***	0.442***	0.477***
	(6.392)	(6.100)	(6.079)	(8.605)	(6.777)	(7.220)	(7.283)	(5.096)	(5.493)
Obs	57	57	57	60	60	60	76	75	75
R-squared	0.107	0.121	0.074	0.217	0.386	0.477	0.083	0.132	0.142

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

gered the interaction between preference and dissonance feeling. For plain flavor, since the risk information was not relevant to it, the premium was reinforced and people who claimed that they preferred the flavor were willing to pay 3 (=$0.186-$0.155) cents more. In contrast, since almond flavor was involved in the food safety information, the preference premium decreased (from 18 cents to 15 cents, then to 13 cents) as participants read the information.

Interestingly, the effect of preference on WTP was negative for the peanut flavor candy bars. This strange relationship can be caused by the trade-off with the treatment effect. The treatment effect (adding premiums to bids due to pre-commitment) was so strong that the preference effect compromised to be negative when fitting the data. Another explanation could be a cross-product effect with plain and almond flavor. As peanut flavor being a more common option in the daily life than the other two alternatives, those who liked it bid for a lower price.

Compared to the significant impact of preference, risk perception seemed to be less influential in affecting WTPs. It almost had no impact on the WTP for plain flavor. For almond, although it was pretty close to be significant in the third round, the magnitude of effect was only as low as 2 cents. Risk perception played a significant role only for peanut flavor in the two informational rounds, yielding 3-cent and 5-cent decrease in WTP for every one-point increase in risk perception. Considering the fact that peanut products were highly involved in the provided food safety information, these significant estimates could be seen as signal for successful manipulation.

Table 18: Study 2_ATE on Changes in Bid - Causal Inference, All

	1	2	3	4	5	6	7	8	9
VARIABLES	dBid_pl21	dBid_pl31	dBid_pl32	dBid_pe21	dBid_pe31	dBid_pe32	dBid_al21	dBid_al31	dBid_al32
Treatment	-0.045	-0.027	0.01	0.081*	0.166***	0.101**	0.059	0.067	0.004
	(-1.177)	(-0.527)	(0.381)	(1.823)	(2.797)	(2.410)	(1.506)	(1.384)	(0.144)
dRisk_21	-0.013			-0.01			-0.046***		
	(-1.018)			(-0.734)			(-3.209)		
dRisk_31		0.009			-0.016			-0.031*	
		(0.513)			(-1.257)			(-1.966)	
dRisk_32			-0.074***			-0.046***			-0.022*
			(-2.920)			(-3.284)			(-1.722)
Risk_2	-0.009			-0.020***			-0.019**		
	(-0.729)			(-2.737)			(-2.214)		
Risk_3		-0.009	-0.003		-0.027**	-0.007		-0.020*	-0.002
		(-0.544)	(-0.312)		(-2.629)	(-1.049)		(-1.945)	(-0.306)
Prefer	0.03	-0.02	-0.043*	0.038	-0.025	-0.069*	-0.025	-0.042	-0.017
	(0.871)	(-0.412)	(-1.741)	(0.955)	(-0.454)	(-1.797)	(-0.642)	(-0.869)	(-0.568)
Constant	0.02	0.037	0.014	-0.031	-0.015	0.012	0.02	-0.02	-0.01
	(0.616)	(0.821)	(0.591)	(-0.654)	(-0.226)	(0.252)	(0.417)	(-0.353)	(-0.276)
Obs	57	57	57	60	60	60	75	75	75
R-squared	0.06	0.028	0.215	0.29	0.289	0.233	0.365	0.263	0.055

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Determinants of Changes in WTP Table 18 showed how changes in WTP were explained by driving factors, i.e. treatment, preference, risk perception and its changes, etc. While preference was more important in explaining pure WTP, risk perceptions had more explanatory power in interpreting the changes in WTP. Note first that dependent variables in Table 18 were changes in WTP from an earlier round to a later round (i.e. $dBid\_pe^{21} = Bid\_pe^{2} - Bid\_pe^{1}$), and hence, were all negative in value. In the table, for peanut and almond flavor, the absolute values of risk perceptions negatively significantly impacted changes in WTP between the last two rounds and the initial round (Column 4, 5 and 7, 8); and changes in risk perception negatively significantly impacted changes in WTP within the last two rounds (Column 6 and 9). On average, 1 point increase in absolute value of risk perception in a certain round (round 2 or 3) yielded about 2 cents' decrease in WTP compared to the first round. 1 point increase in the changes of risk perception yielded about 4.6 cents' decrease in WTP for peanut between second and third round and 2.2 cents' decrease for almond. Preference, however, was only significant to changes between the last two rounds for plain and peanut flavors.

In addition to the impact of preference and risk perceptions, treatment still played its role in mitigating changes in WTPs, significant and large (in magnitude) for peanut flavor, and close to significant and relatively smaller (in magnitude) for almond flavor. These positive estimates of the

Table 19: Study 2_ATE on Changes in Bid - Causal Inference, Preferred Only

	1	2	3	4	5	6	7	8	9
VARIABLES	dBid_pl21	dBid_pl31	dBid_pl32	dBid_pe21	dBid_pe31	dBid_pe32	dBid_al21	dBid_al31	dBid_al32
Treatment	-0.047	-0.023	0.011	0.124**	0.179***	0.052	0.067	0.045	-0.033
	(-0.915)	(-0.347)	(0.407)	(2.726)	(3.327)	(1.502)	(1.348)	(0.837)	(-1.242)
dRisk_21	-0.02			-0.015			-0.045**		
	(-1.096)			(-0.763)			(-2.446)		
dRisk_31		0.014			-0.013			-0.023	
		(0.548)			(-0.877)			(-1.387)	
dRisk_32			-0.110***			-0.037**			-0.015
			(-3.755)			(-2.677)			(-1.398)
Risk_2	-0.01			-0.022**			-0.022**		
	(-0.402)			(-2.325)			(-2.104)		
Risk_3		-0.002	0.005		-0.040***	-0.013		-0.025**	0
		(-0.0673)	(0.422)		(-3.083)	(-1.555)		(-2.216)	(-0.0453)
Constant	0.052	0.003	-0.045	-0.022	0.038	0.031	0.002	-0.022	0.008
	(0.785)	(0.035)	(-1.265)	(-0.286)	(0.414)	(0.519)	(0.035)	(-0.317)	(0.224)
Observation	33	33	33	35	35	35	50	50	50
R-squared	0.081	0.014	0.347	0.391	0.448	0.331	0.387	0.295	0.071

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

treatment effect were evidence of sticky behaviors due to consumers' taste of consistency.

Table 19 showed the robustness check for the above findings by using only those who preferred the certain flavor. Comparing Table 19 with Table 18, one can find that the main results are consistent.

3.3.7 Market Demand and Price Elasticity

Individual responses to food safety information under different treatment situations were discussed in the above sections. In this section, the focus is the aggregated results derived from each group. This section provides insights about market responses to food safety information for different groups of customers.

In order to investigate market responses, aggregate demands for each flavor were derived from the following subgroups: 1) control group, 2) treatment group, 3) control group with only subjects who preferred the flavor, 4) control group with only subjects who did not prefer the flavor and 5) all subjects in both control and treatment groups. The aggregate demands were then converted to demand shares for each subgroup so as to make all 5 sub-groups comparable.

Table 20 listed the regression results of the fitted inverse demand functions for each flavor-subgroup combination. Equation 5 is the inverse demand functional form used in fitting regressions in Table 20.

Table 20: Study 2_Inverse Demand Curve by Group

	Plain			Peanut					Almond				
	Control	Treat	ctrl prefer	Control	Treat	ctrl prefer	ctrl no-pref	all	Control	Treat	ctrl prefer	ctrl no-pref	all
VARIABLES	price	price	price	price	price	price	price	price	price	price	price	price	price
Demand	-0.010***	-0.009***	-0.009***	-0.009***	-0.008***	-0.008***	-0.010***	-0.010***	-0.010***	-0.010***	-0.009***	-0.010***	-0.010***
	(-25.16)	(-12.94)	(-16.00)	(-22.93)	(-8.665)	(-14.34)	(-26.29)	(-30.66)	(-27.18)	(-26.87)	(-16.29)	(-25.13)	(-29.47)
Round2	0.015	-0.009	0.034	-0.133***	-0.03	-0.075	-0.151***	-0.120***	-0.153***	-0.070**	-0.155***	-0.144***	-0.124***
	(0.511)	(-0.170)	(0.748)	(-3.991)	(-0.390)	(-1.469)	(-5.126)	(-4.754)	(-5.354)	(-2.478)	(-3.367)	(-4.683)	(-4.725)
Round3	0.007	-0.024	-0.006	-0.195***	-0.067	-0.149***	-0.207***	-0.180***	-0.176***	-0.091***	-0.187***	-0.160***	-0.140***
	(0.217)	(-0.427)	(-0.126)	(-5.736)	(-0.874)	(-2.892)	(-6.931)	(-7.054)	(-6.122)	(-3.213)	(-4.003)	(-5.202)	(-5.333)
Constant	1.076***	1.168***	1.125***	1.054***	1.152***	0.898***	1.098***	1.121***	1.131***	1.208***	1.205***	1.054***	1.175***
	(34.420)	(18.040)	(22.150)	(31.430)	(12.460)	(19.840)	(35.930)	(41.780)	(37.090)	(36.520)	(22.400)	(34.370)	(40.100)
Obs	33	33	33	33	33	33	33	33	33	33	33	33	33
R-squared	0.956	0.852	0.898	0.948	0.721	0.876	0.96	0.97	0.962	0.961	0.901	0.956	0.968

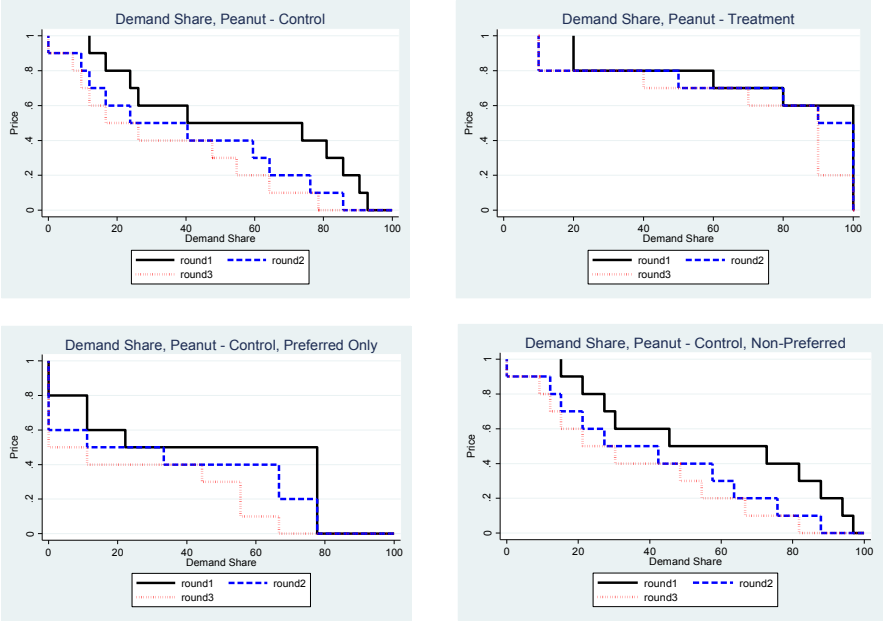
t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

$$Price_{(flavor,\,group)} = a + bDemand_{(flavor,\,group)} + \sum_{k=2}^{3} \mu_k + \epsilon_{(flavor,\,group)} \tag{5}$$

In Equation 5, the inverse demand function was specific to "flavor, group" combination. There were 3 flavors and 5 subgroups of subjects, each representing a customer pool in the real market. Price was exogenously given by the design of the study. 11 price values ranged from \$0 to \$1, with increment of 10 cents. Demand was calculated as the inverse cumulative percentage in accordance to each price value (i.e. the demand D at a certain price P is the percentage of the biddings that are higher than or equal to P). The constant a represented the intercept of the inverse demand for the first round (default, no information). b was the slope effect. By the construction of the model, price elasticity can be calculated by $e = \frac{dDemand/Demand}{dPrice/Price}$. μ_k represented the shift of the inverse demand curve caused by the information in either round 2 or round 3. Focus was given to the comparison of these estimates across subgroups so as to capture different market responses to risk information.

According to the results in Table 20, the slopes of the demand curves were roughly 0.1 for all cases. In response to risk information, demand for peanut and almond flavor both shifted downward, while demand for plain flavor did not, which once again suggested the successful manipulation of risk information. However, for both peanut and almond flavor, the shift of demand curve was not the same across different subgroups. Take peanut flavor for example, the demand curve shifted 13 cents downwards in the control group (Column 4), which meant at any given demand level (in terms of market demand share), the price was 13 cents less in round 2 than in round 1. In contrast, the treatment group did not show any significant downward shift after the same risk information

Figure 2: Study 2_Demand Curves in Percentage Share - Peanut



being revealed to participants (Column 5). Similarly, in the third round, the downward shift was 20 cents for control group, significant at 1% level; and only 7 cents for the treatment group, but statistically insignificant. Differences remained when comparison was made between treatment and control with only those who preferred peanut flavor (Column 6). Same for almond flavor, even though the demand curve significantly shifted down for both treatment (Column 10) and control group (Column 9), the magnitude of shift was always larger in control group than in treatment (15 cents vs. 7 in round 2 and 18 cents vs. 9 in round 3).

Figure 2 to 4 showed the shift of demand curves across rounds for each flavor and subgroup. Differences in the shift of demand curve suggested that customers with long-last preference or pre-commitment were less responsive to food safety information regarding the product they previously chosen. Psychological biases were crucial in influencing the efficacy of public health communication with consumers.

3.4 Discussion

This study investigated how individual consumers react to food safety information and make purchase decision. Using an incentive compatible auction mechanism, this study elicited consumers' WTP under different informational settings. Consistent to the findings in psychology, consumer's

Figure 3: Study 2_Demand Curves in Percentage Share - Almond

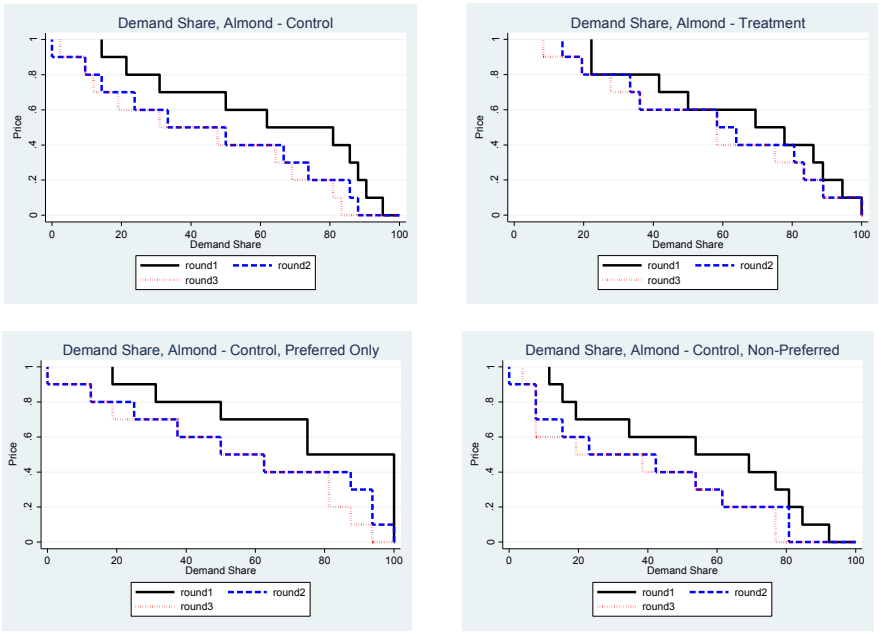
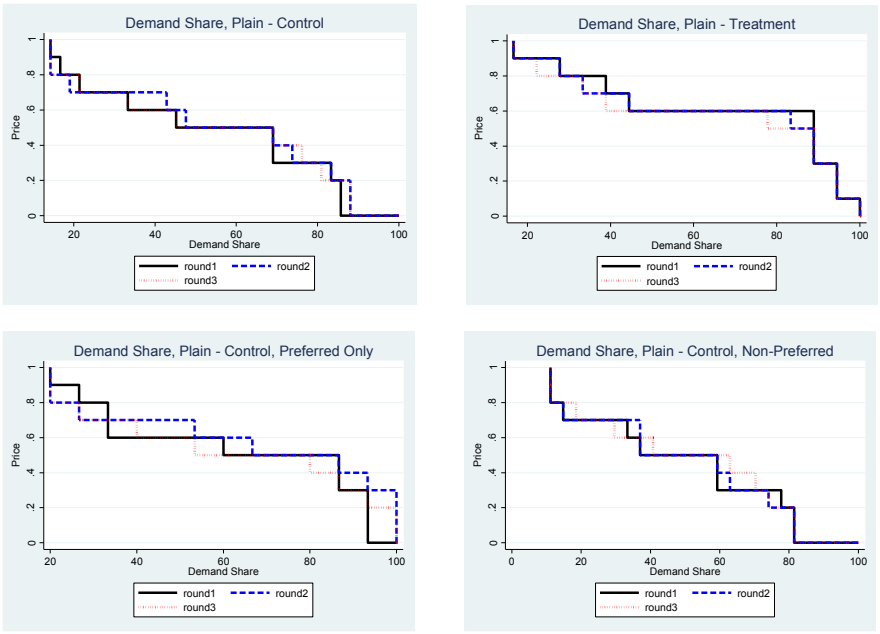


Figure 4: Study 2_Demand Curves in Percentage Share - Plain



judgment and information processing depend a lot on their initial beliefs or consumption status.

Results showed that consumers bid much higher when they freely chose food items (treatment) than when they were randomly assigned (control), suggesting cognitive dissonance. On average, the bidding premium was about 13 cents (roughly 30%) higher for low-risk food item and 30 cents (almost 60%) higher for high-risk item. The bidding premiums were further enlarged as food safety information was revealed to consumers. Confirmation bias hypothesis was supported by the finding that free-choice group was more reluctant to change the bids despite of increased risk perceptions. In terms of market responses, due to psychological biases among consumers, demand curves were less possible to shift down under food safety risk.

Results in this study suggested that consumers were less responsive to public information due to their existing habits. Extra strategies would be needed to increase the efficiency of public communication to promote health.

References

[1] Akerlof, George A & Dickens, William T. 1982. "The Economic Consequences of Cognitive Dissonance," American Economic Review, 72(3): 307-319.

[2] Doob, A. N., J. M. Carlsmith, J Freedman, T. K. Landauer and S. Tom. 1969. "Effect of Initial Selling Price on Subsequent Sales", Journal of Personality and Social Psychology, 11 (April): 345-350.

[3] Downs, Julie S., George Loewenstein, and Jessica Wisdom. 2009. "Strategies for Promoting Healthier Food Choices." American Economic Review, 99(2): 159–64.

[4] Ehrlich, D., I. Guttman, P. Schonbach and J. Mills. 1957. "Post-decision Exposure to Relevant Information", Journal of Abnormal and Social Psychology, 54 (Jan): 98-102.

[5] Engel, J. F. 1963. "Are Automobile Purchases Dissonant Consumers?" Journal of Marketing, 27 (April): 55-58.

[6] Festinger, L. 1957. A Theory of Cognitive Dissonance, Stanford, CA: Stanford.

[7] Fox, J. A., B. L. Buhr, J. F. Shogren, J. B. Kliebenstein and D. J. Hayes. 1995. "A Comparison of Preferences for Pork Sandwiches Produced from Animals with and without Somatotropin Administration", Journal of Animal Science, 73: 1048-1054.

[8] Freedman, J. L. and D. O. Sears. 1965. "Selective Exposure", in L. Berkowitz, Advances in Experimental Social Psychology, Vol. 2 New York: Academic Press.

[9] Frey, D. 1986. "Recent research on selective exposure to information", in L. Berkowitz (Ed.), Advances in experimental social psychology, 19: 41–80. New York: Academic Press.

[10] Hayes, D. J, J. F. Shogren, S. U. Shin and J. B. Kliebenstein. 1995. "Valuing Food Safety in Experimental Auction Markets", American Journal of Agricultural Economics, 77L 40-53.

[11] Kassajian, H. H. and J. B. Cohen. 1965. "Cognitive Dissonance and Consumer Behavior", California Management Review, 8 (Fall): 55-64.

[12] Lin, C. J., J. Lee and S. T. Yen. 2004. "Do Dietary Intakes Affect Search of Nutrient Information on Food Labels?", Social Science and Medicine, 59: 1955-1967.

[13] LoSciuto, L. and R. Perloff. 1967. "Influence of Product Preference on Dissonance Reduction", Journal of Marketing Research, 4 (March): 286-290.

[14] Lusk, J. L., J. A. Fox, T. C. Schroeder, J. Mintert and M. Koohmaraie. 2001. "In-Store Valuation of Steak Tenderness", American Journal of Agricultural Economics, 83: 539-550.

[15] Lusk, Jayson L., L. O. House, C. Valli, S. R. Jaeger, M. Moore, B. Morrow and W. B. Traill. 2004a. "Effect of Information about Benefits of Biotechnology on Consumer Acceptance of Genetically Modified Food: Evidence from Experimental Auctions in United States, England, and France", European Review of Agricultural Economics, 31(July): 179-204.

[16] Lusk, J. L., T. Feldkamp and T. C. Schroeder. 2004b. "Experimental Auction Procedure: Impact on Valuation of Quality Differentiated Goods", American Journal of Agricultural Economics, 86: 389-405.

[17] Lusk, J. L., M. Rousu. 2006. "Market Price Endogeneity and Accuracy of Value Elicitation Mechanisms", Using Experimental Methods in Environmental and Resource Economics, John A. List (ed.) Northhampton, MA: Edward Elgar Publishing.

[18] Noussair, C., S. Robin and B. Ruffieux. 2004. "Revealing Consumers' Willingness-to-Pay: A Comparison of the BDM Mechanism and the Vickrey Auction", Journal of Economic Psychology, 25: 725-741.

[19] Rutstrom, E. E. 1998. "Home-Grown Values and Incentive Compatible Auction Design", International Journal of Game Theory, 27: 427-441.

[20] Rousu, M., D. C. Monchuk, J. F. Shogren and K. M. Kosa. 2005. "Consumer Willingness to Pay for 'Second-Gerneration' Genetically Engineered Products and the Role of Marketing Information", Journal of Agricultural and Applied Economics, 37: 647-657.

[21] Shogren, J. F., M. Margolis, C. Koo and J. A. List. 2001. "A Random nth-Price Auction", Journal of Economics Behavior and Organization, 46: 409-421.

[22] Shogren, Jason F. and Tommy Stamland. 2007. "Valuing Lives Saved from Safer Food: A Cautionary Tale Revisited", American Journal of Agricultural Economics, 89: 1176-1182.

[23] Turvey, Calum. 2008. "Risk, Fear, Bird Flu and Terrorist: A Study of Risk Perceptions and Economics", working paper, Cornell University.

[24] Wessells, C. R., J. Kline and J. G. Anderson. 1996. "Seafood Safety Perceptions and Their Effects on Anticipated Consumption under Varying Information Treatments", Agricultural and Resource Economics Review, (April): 12-21.

4 Chapter 4

Study 3: Credit Rationing under Lenders' Risk-Aversion and Asymmetric Adverse Incentives

4.1 Introduction

Incentive mechanisms are important attributes to financial decision-making. The incentives influence lenders' loan decisions, which in turn affect the profitability of financial institutes and the availability of credit supply in the market. These impacts have far-reaching implications. For example, accessibility to micro credit is crucial for economic development in rural and poor areas. But if correct incentives are not in place, lenders may ignore the social significance of their decisions, and avoid lending to potentially high risk micro credit clients. The tranditional approach to understanding the lender-borrower relationship has been brought to economics of credit rationing[1] (Stiglitz and Weiss, 1981). However, some recent research began to link credit rationing to incentive mechanisms (Banerjee and Duflo, 2008). Under this reasoning, the standard credit rationing model may not be a complete explanation; rather in an environment with incentives a form of policy institutional rationing is exposed, in which it is the rules of the game that lead to rationing, and not borrowers' risk per se. This study complements the literature along this stream by showing that certain incentive schemes can result in credit rationing through lenders' risk-aversion and behavioral responses. It empirically extends the incentive mechanism design in banking to a behavioral scope. Furthermore, the results contribute to the increase of profitability in financial institutes, help to alleviate credit rationing and stabilize credit supply in the market.

The study is set in the micro credit loan market for farm households in China. Rural Credit Cooperatives (RCCs) provide the majority of financial services to farmers in rural China. In order to control credit risk and maintain profitability, the CBRC (China Banking Regulatory Commission) has provided guidelines to RCCs to adopt a performance-based internal incentive structure for their loan officers, called the Personal Responsibility System (PRS). The PRS provides bonuses to loan officers for approved loans that are in performance (or performing loans, PLs) while imposing penalties on non-performing loans (NPLs). While this incentive structure successfully keeps the

[1] By Stiglitz and Weiss (1981), credit rationing happens if some applicants receive loans, while others with similar credit risk do not even when they are willing to offer higher interest rates.

NPL ratio low, it may have the adverse incentive of encouraging lenders to reject loans that would have otherwise been good. This adverse incentive not only threatens the profitability of RCCs, but also results in unnecessary credit rationing of farmers[2]. So far, this incentive mechanism in China has never been questioned, let alone tested. This study investigates how incentive structures like the PRS affect the lending practice of loan officers.

How agents respond to incentives is difficult to quantify. In this study, experimental techniques were developed to investigate the incentives faced by front-line rural lenders under China's Personal Responsibility System (PRS). In the experiment, loan officers from local RCCs in Shandong, China were recruited to evaluate randomly selected loan applications and make lending decisions. All loan files were previously approved with known performance and repayment status. Each loan officer was randomly assigned to one of two incentive groups. One was analogous to pure PRS, and the other was the PRS with an additional dis-incentive for Type II error (i.e. lose monetary payoff when rejecting good loans). For both groups, the expected monetary payoff was roughly one day salary. The two groups were further randomized over prior knowledge about the probability distribution of the application pool. Half of them were informed that loan files were randomly selected from a database with 50:50 PLs and NPLs, while the other half were not informed.

This experimental approach has the following important features that can help uncover the relationship between incentive mechanisms and lending behavior. First, a control group with incentive structure proportionally mimicking the PRS generates a benchmark to better understand the lending behaviors. Second, exogenously varying incentive structure by adding an additional dis-incentive (penalty) for Type II error induces clear estimation of the relative strengths and/or weaknesses of managerial incentives. Third, randomly introducing probability knowledge about the application pool reveals the interdependent effects of prior perception and incentive schemes on decision-making. Fourth, using previously approved loan files makes decision outcomes fully observable and comparable. Hence, incentive efficacy, profitability and credit rationing can be examined under alternative institutional settings.

[2]High frequency of loan misclassification affects the stability and profitability of financial institutes. For Type I error (i.e. approve a bad loan), financial institutes lose the principal, and for Type II error (i.e. reject a good loan), financial institutes forgone the revenues associated with good loans. As Nayak and Turvey (1997) argued, with excess demand of funds, a lender will not keep money idle. So the cost of Type II error can be considered as the difference between the revenue forgone from a good borrower and the expected profit from an alternative loan. When the alternative loan is of high risk, Type II error could be rather costly. As a key financial institute for microcredit loans, profitability of RCCs immediately influences the stability and sustainability of financial supply to farmers.

The efficacy of incentive structures can be estimated by the mean differences between treatment conditions. In the *Baseline Condition (BC)*, the incentive structure was set proportional to pure PRS without revealing probabilities about the application pool. In the *Dis-incentive Condition (DC)*, loan officers were given an additional dis-incentive (penalty) for Type II error (i.e. rejecting loans that are in performance) when they were making lending decisions. Probability information was not offered in this condition, either. In the *Probability Condition (PC)*, the loan officers kept using pure PRS, but were offered probability distribution information about loan applications. In the *Mixed Condition (MC)*, both dis-incentive for Type II error and prior probability knowledge were imposed on the loan officers. This counterfactual model will reveal the degree of risk avoidance brought about by adverse incentives by examining differences in patterns of loan acceptance and rejection with Type I (i.e. approving non-performing loans) and Type II error. Combining decisions with specific loan file features, aggregate loan supply changes in the market, and profitability changes in the RCCs can then be assessed. In addition, the interdependence between prior probabilities and incentives can be analyzed through the experimental conditions in place.

Results showed that in the Baseline Condition, loan officers on average approved 44% of the applications. Due to a sense of risk and/or loss aversion, they were 12.9% less likely to approve a loan under pure PRS than when there was additional dis-incentive for Type II error, suggesting credit rationing. As a result, the frequency of Type II error was 10.2% higher under PRS, which implied up to 41.52% decrease of interest returns in RCCs and 37.7% decrease of credit supply in the financial market. Moreover, ambiguity about riskiness of the application pool made credit rationing even more severe. Providing probability information removed ambiguity and offset the negative prior. Approval rates increased about 12% with probability information. This yielded a significant decline of Type II error at the cost of an insignificant jump in Type I error. This is a significant result for not only does it reveal that adverse incentives can affect loan decisions, but also ambiguity about loan riskiness acts along with the incentive mechanisms to affect loan outcomes.

Evidence from the study generates the following implications. First, the PRS forces a loan officer to face the risk of penalty in case of NPLs, but at the same time implicitly offers them a "secure" option of rejecting any loans to avoid the risk. As a result, a risk-averse loan officer chooses the "secure" option to "insure" the risk of loss. This conservative behavior generates credit rationing in the loan market, increases Type II error and lowers interest returns of RCCs. Second, when

balanced (Type I and Type II) incentives are in place, loan officers are relatively less averse to Type I error, because they must now contend with Type II error and balance. Finally, providing prior information about the application pool offsets ambiguity and leads to decision accuracy.

Due to the challenge of observing internal dynamics of a bank's managerial processes, and measuring their impacts on decision-making and lending, empirical evidence about institutional efficacy in banking remains scarce. Results of this research complement the literature in the following fields: theories of credit rationing (Stiglitz and Weiss, 1981; Jaffee and Russell, 1976; and Leland and Pyle, 1977); credit rationing in developing countries (Ghosh et al., 2000; Karlan and Morduch, 2009; Banerjee, Cole and Duflo, 2008); empirical literature on agency problems in banks (Liberti 2003, Liberti and Mian 2009, Agarwal and Wang 2009, Hertzberg et al. 2010); incentives within firm (Lazear, 2000; Bandiera, Barankay, and Rasul, 2007, 2009; Paarsch and Shearer, 2009; Bandiera, Barankay, and Rasul, 2010); bank function and organizational design (Berger et al., 2001; Berger and Udell, 2002; Petersen and Rajan, 2002; Berger et al., 2005; Mian, 2006); and field experiments about risk-taking and credit decisions in a micro finance context (Gine, Jakiela, Karlan, and Morduch, 2010; Fischer, 2010; Kanz, 2010).

Theoretically, the study extends the principal-agent problem and incentive mechanism design in banking to a behavioral scope. In practice, this study contributes to the increase of stability and profitability of financial institutes in general and RCCs in particular. For banking policy, this study adds value to credit rationing alleviation and loan supply stabilization in the financial market. More profoundly, this study imposes positive effects on sustainable accessibility to financial services in less developing areas and helps to promote social development and poverty alleviation in the long-run.

The remainder of the paper proceeds as follows. Section 4.2 reviews related literature. Section 4.3 describes the background of micro credit market and institutional arrangements in China upon which the experiment was set. Section 4.4 provides a behavioral model of loan decision-making to guide the design of experiment. Section 4.5 outlines the experimental design and method of randomization. Section 4.6 presents the empirical results. Section 4.7 discusses robustness and validity. Section 4.8 discusses and concludes the paper.

4.2 Literature Review

Ghosh et al. in their review of credit rationing draw reader's attention to two different forms of quantity constraints: *micro* credit rationing, which places credit limits on borrowers (below first-best levels), and *macro* credit rationing, which randomly denies access to any credit to a fraction of borrowers. The second form involves asymmetric treatment of otherwise identical agents. They showed that both forms of rationing might coexist, and play complementary but distinct roles. The second form of rationing gains in importance when information flow within the lending community is poor.

In our study, in the Chinese context, the PRS is dictated as central policy by the CBRC and is aligned with Ghosh et al.'s *macro* view. In this context, we identify these observations about *macro* effects with what we have referred to as "policy rationing". The separation of these policy impacts from Stiglitz and Weiss's credit rationing at the lender-borrower level is in fact a useful departure, which permits one to review a number of related studies in a broader context.

Following the stream of theoretical approaches, many studies devoted in efficient mechanism design and improving institutional settings. There are two major directions. One is on borrowers' side and the other is on lenders side.

Literature on borrower incentives has many interesting studies. For example, Fischer (2008) developed a theory that unifies models of investment choice, informal insurance, and formal financial contracts. He then tested model predictions using a series of experiments with clients of a large microfinance institution in India. The experiments confirmed that joint liability creates two incentive efficiencies. First, borrowers free-ride on their partners, making risky investments without compensating partners for this risk. Second, the addition of peer-monitoring overcompensates, leading to sharp reductions in risk-taking and profitability.

Gine, Jakiela, Karlan, and Morduch (2010) systematically unpacked microfinance mechanisms through ten experimental games played in an experimental economics laboratory in urban Peru. They found dynamic incentives strongly reduce risk-taking even without group-based mechanisms. Group lending increases risk-taking, especially for risk-averse borrowers, but this is moderated when borrowers form their own groups. Group contracts benefit borrowers by creating implicit insurance against investment losses, but the costs are borne by other borrowers, especially the most risk averse.

Cassar and Wydick (2008) addressed the effect of social capital on group lending by conducting an artifactual experiment in five countries. They carried out treatments for social homogeneity, group monitoring, and self-selection. Results showed contribution rates differ substantially between countries, and the influence of different types of social capital varies depending on context, that group lending appears to create as well as harness social capital, and that peer monitoring can have perverse as well as beneficial effects on group performance. They also distinguished between spiritual capital and social capital among religions, finding mild evidence for the effect of spiritual capital on borrowing group performance.

Our study focuses primarily on lender incentives. Most relevant research includs but is not limited to, for example, Banerjee, Cole and Duflo (2009) who used observational data in India to empirically test a model of lending in which loan officers face both an incentive to lend, and the possibility of penalties for making loans that go bad. They found evidence that following the discovery of a fraud in a particular bank branch, vigilance activities greatly increases. This in turn results in reduced lending: the amount of credit declines sharply at the affected bank branch, as well as neighboring branches. This effect is large, and persists in part for up to two years. Bank risk-taking also declines following an inspection.

Cull, Demirgüç-Kunt and Morduch (2009) found empirical evidence that regulation and supervision negatively impact profitability of financial institution, and are associated with substantially large average loan sizes and less lending to women. The pattern is consistent with the notion that profit-oriented microfinance institutions absorb the cost of supervision by curtailing outreach to market segments that tend to be more costly per dollar lent. By contrast, microfinance institutions that rely on non-commercial sources of funding (for example, donations), and thus are less profit-oriented, do not adjust loan sizes or lend less to women when supervised.

Kanz (2010) analyzed the effect of organizational structure on bank lending by using a framed field experiment in Indian market for small enterprise loans. Results showed that supervision reduces defaults and increases loan-level profit, but at the same time discourages collection and use of qualitative information. Incentive contracts such as performance pay can moderate the adverse effects. Findings shed new light on the nature and importance of agency conflict within the bank, and suggest that performance pay can play an important role in mitigating information and agency problems in the provision of entrepreneurial finance in an emerging market.

Our study also falls into broader empirical literature on agency problems in banks, incentives within firm, and bank function and organizational design. Typical studies are, for example: Hertzberg et al. (2010) who presented evidence that reassigning tasks among agents can alleviate moral hazard in communication. Based on a commercial banks internal reporting data they showed that agents do not report bad news if it reflects poorly on their own ability. A rotation policy that routinely reassigns loan officers to borrowers changes this reporting behavior. When an officer anticipates rotation, reports are more accurate and contain more bad news about the borrowers repayment prospects. Further, improved communication has first order effects on lending outcomes.

Agarwal and Wang (2009) used a unique data set on small business loan officer compensation from a major commercial bank to test model predictions that incentive compensation increases loan origination, but may induce loan officers to book more risky loans. They found that the incentive package amounts to a 47% increase in loan approval rate, and a 24% increase in default rate. Overall, the bank loses money by switching to incentive pay.

For incentives within firms, Bandiera, Barankay, and Rasul (2010) presented evidence from a field experiment designed to evaluate the impact of rank incentives and tournaments on the productivity and composition of teams. Strengthening incentives, either through rankings or tournaments, makes workers more likely to form teams with others of similar ability instead of with their friends. Introducing rank incentives however reduces average productivity by 14%, whereas introducing a tournament increases it by 24%. Results implied that provision of team-based incentives crowds out the productivity enhancing effect of social connections under team production.

For bank function and organizational design, Mian (2006) showed that greater cultural and geographical distance between a foreign bank's headquarters and local branches leads it to further avoid lending to "informationally difficult" yet fundamentally sound firms requiring relational contracting. Greater distance also makes them less likely to bilaterally renegotiate, and less successful at recovering defaults. These distance constraints can be large enough to permanently exclude certain sectors of the economy from financing by foreign banks.

4.3 Incentive Mechanisms and the Personal Responsibility System

Rural Credit Cooperatives (RCCs) have been identified as a key vehicle for the delivery of financial services to small-scale entrepreneur and customers in rural areas in China. RCCs were first

established during the rural corporative movement in 1950s. They account for 11.5% of deposits (CNY[3] 2,233 billion), 10% of loans outstanding (CNY 1,618 billion) of the banking sector, and 85% of agriculture loans. China has 32,397 RCCs with about 628,000 employees, 2,441 RCC unions (RCCUs) at the county level, 65 RCCUs at the prefecture level, and 6 RCCUs at provincial level[4]. Agriculture loans outstanding by RCCs in June 2003 amounted to CNY 700 billion. Farmer loans amounted to CNY 555 billion and microcredit loans amounted to CNY 114 billion.

Though RCCs play an important role in providing financial services in rural areas, they still have many institutional problems, among which historically high frequency of non-performing loans (NPLs) was a biggest one. Nowadays, in order to avoid approving loans that are of high credit risk (loans which have high potential to become NPLs), while at the same time still keep promoting loan amount to maintain profit level, RCCs adopt an incentive mechanism called the Personal Responsibility System (PRS). Briefly speaking, the PRS is an internal institution of credit risk management in RCCs. The PRS on the one hand, provides bonuses to loan officers for approved loans that are in performance; and on the other hand, imposes penalties to those who are responsible for NPLs in case of loss.

To be more detailed, PRS first specifies responsibility of loan application shared by relevant personnel. In a typical microcredit loan application procedure, the primary loan officer meets the loan applicant, investigates the household and/or business situation on-site, writes credit reports and finally put together the application package. A secondary loan officer reviews application materials and makes a joint decision of approving or rejecting with the primary officer. Two managers supervise the whole procedure and provide guidance when necessary. Upon approval, the primary loan officer takes responsibility of 60% of the loan amount. The secondary loan officer takes another 30% and two managers take 5% each.

During the loans outstanding period, in addition to the base salary, loan officers also obtain performance salary based upon the performance of loans they are responsible for. The monthly performance salary comes from two parts: first, a loan officer gets CNY 0.5/10,000 of performing loan amount; second, he gets another CNY 300/10,000 from the interests generated from performing loans. If there is no interest or a repayment delay, 90% of the total monthly performance salary is

[3]Chinese Yuan

[4]Statistics as of 2005.

given to loan officers as bonus and 10% is withheld as risk deposit. The 10% deposit will then be returned to loan officers in the first month of next year. In case of interest or a repayment delay, a loan officer gets only CNY 0.2/10,000 of loan amount from NPLs and CNY 150/10,000 of delayed interests collected later. Moreover, the risk deposit under his account will be deducted to (partially) recover the losses. A loan officer is also responsible for the collection of delayed repayments. If he cannot finish his monthly assignment in full, but can finish more than 50%, he gets a proportion of his monthly performance salary. The proportion equals to the percentage of the monthly assignment fulfilled. If, instead he finishes less than 50% of the assignment, he loses all the monthly performance salary. In the RCCs in Shandong Province which collaborated in our experiments, a loan officer on average gets CNY 2,000-3,000 per month for base salary and CNY 4,000 to 5,000 for performance salary. In case of delayed payments and/or NPLs, a loan officer can lose up to all of his performance salary, which is roughly two times of the base salary, or 70% of his total monthly income.

The Personal Responsibility System (PRS) did a good job in keeping the NPL ratio low. Within the past 10 years, together with many other reforms, the PRS has decreased the NPL ratio in RCCs sharply from 44% to 5%[5]. However, performance metrics such as these may in fact be overestimating performance because they exclude unobserved losses from good loans that were ultimately rejected. The proposition guiding this research is that under the PRS, when there is a chance of being penalized, a risk-averse loan officer can be more inclined to reject a loan than a risk-neutral one so as to avoid the risk of penalty. When the penalty can make a loan officer lose what he already owned, a sense of loss aversion can further reinforce the tendency to reject any loans. While the institutional setting of the PRS keeps the frequency of Type I error being low, a potential side-effect caused by loan officers' decision behaviors is a high frequency of Type II error.

4.4 Theoretical Model

In order to better understand loan decision behaviors associated with credit risk and make a clear guidance for the experimental design, a behavioral model is proposed in this section to describe the decision procedure. It is assumed that for any lending decision a loan officer makes, he expects to experience a standard "monetary utility" [6] and a "gain-loss utility". In deterministic environ-

[5]"Rural Credit Cooperatives in China" in *Planet Finance*, p1-4.

[6]It is the so-called "intrinsic utility" under the classic utility theory framework. Sometimes, it is also called "consumption utility". Since a loan officer never "consumes" a loan, but rather makes decision of either approving or

ments, a loan officer choices to maximize monetary utility, but gain-loss utility influences behaviors when uncertainty about credit risk gets involved. The model equates the reference point with the probabilistic beliefs about outcomes and takes it as endogenous in the decision environment. Personal equilibrium requires a loan officer correctly predicts the probabilistic environment and makes consistent optimal behavior given the correct expectation.

The model shows that under the Personal Responsibility System (PRS) where there is no penalty for Type II error of loan misclassification (i.e. rejecting good loans), a loan officer behaves first-order risk aversion and is inclined to reject a loan to avoid potential penalty for non-performing loans (NPLs), suggesting credit rationing. However, when an additional penalty for Type II error is in place, a loan officer, who now anticipates a full set of risk for penalty of any misclassification (both Type I and Type II error), balances the two, decreases aversion to Type I error (i.e. approving bad loans) and becomes more inclined to approve a loan than before. As a result, the credit rationing problem is alleviated. Further, combining the model with dis-proportional probability weighting, it suggests that ambiguity towards the application pool worsens the prior perception and lowers the approval rates.

4.4.1 Incentive Structures

This section discusses the incentive structures loan officers face when making decisions.

A loan officer faces a lottery $F_a^t = \left(\theta,\, B_a^t,\, P_a^t\right)$ if he approves a loan and $F_r^t = \left(\theta,\, P_r^t,\, B_r^t\right)$ if he rejects. The superscript $t = 0,\, 1$ represents the treatment status, 0 if it is basic PRS with no other adjustment, 1 if there is an additional penalty for rejecting good loans (i.e. Type II error). The subscripts $d = a,\, r$ represents the loan decision, approve and reject respectively. $\theta \in [0,\, 1]$ is the perceived probability of a loan being a good one[7]. $B_d^t \geq 0$ and $P_d^t \leq 0$, $d = a,\, r$ are bonus and penalty under corresponding decision cases[8].

The lottery $F_a^t = \left(\theta,\, B_a^t,\, P_a^t\right)$ can be interpreted as when approving a loan, a loan officer gets bonus B_a^t with probability θ if the loan performs well and gets penalty P_a^0 with probability $(1 - \theta)$ if

rejecting it and gets monetary incentives according to the loan performance later, we call it "monetary utility".

[7] Or the perceived probability of repayment.

[8] Though there is a complete system to calculate bonus and penalty according to the loan amount, interest rates, payment delayed periods, etc, for simplicity concern, fixed bonus and penalty schemes are used here and in the remaining part of the model. The properties and predictions will survive with variant incentives merely by replacing the fixed incentives with functions of all relevant factors, such as $B_i = B_i\left(\cdot\right)$ and $P_i = P_i\left(\cdot\right)$.

it fails to perform well. Similarly, the lottery $F_r^t = \left(\theta,\, P_r^t,\, B_r^t\right)$ can be interpreted as when rejecting a loan, a loan officer gets penalty P_r^t with probability θ if the loan actually performs well (penalty for Type II Error) and gets bonus B_r^t with probability $(1-\theta)$ if it actually fails to perform well.

Under the Personal Responsibility System (PRS), a loan officer gets bonuses for approved loans that are in good performance and gets penalties for non-performing ones. Since the performance status of rejected loans is unobservable, there is no payment adjustment for those, i.e. $P_r^0 = B_r^0 = 0$, and hence, the lottery of rejecting a loan shrinks to a constant $F_r^0 = 0$. In order to explore potential effects of the PRS, an additional penalty for Type II error is added to the basic PRS structure. Use superscript 1 to indicate the treatment of additional penalty. Following the definition above, when a loan officer approves a loan, he faces a lottery $F_a^1 = \left(\theta,\, B_a^1,\, P_a^1\right)$; when rejecting a loan, he faces a lottery $F_r^1 = \left(\theta,\, P_r^1,\, B_r^1\right)$. To make the treatment comparable to the basic PRS, let $F_a^1 = F_a^0 = F_a = (\theta,\, B_a,\, P_a)$, $B_r^1 = B_r^0 = 0$ and $P_r^1 < 0$, that is $F_r^1 = \left(\theta,\, P_r^1,\, 0\right)$ but $F_r^0 = (\theta,\, 0,\, 0) = 0$. $P_r^1 < 0$ represents a dis-incentive treatment for Type II Error, which is the only difference from the PRS.

When a loan officer makes decision fully depending on the perceived probability θ, he is facing a compound lottery of $F^t = \left(f\left(\theta\right),\, F_a^t,\, F_r^t\right)$. $f\left(\cdot\right) \in [0,1]$ is a probability distribution. The compound lottery indicates a mixed strategy for a loan decision, that is approving a loan with probability $f\left(\theta\right)$ and facing the lottery F_a^t; and rejecting a loan with probability $1-f\left(\theta\right)$ and facing the lottery F_r^t. The compound lotteries under two incentive schemes are now $F^0 = \left(f\left(\theta\right),\, F_a,\, 0\right)$ and $F^1 = \left(f\left(\theta\right),\, F_a,\, F_r^1\right)$. Under different incentive schemes, a representative officer optimally chooses different $f\left(\cdot\right)$ [9] to guide his decision.

4.4.2 Reference-Dependent Utility

A reference-dependent utility $u\left(w \,|\, r\right) = m\left(w\right) + \mu\left(m\left(w\right) - m\left(r\right)\right)$ is adopted to capture a loan officer's feeling when making loan decisions under a riskless wealth outcome $w \in R$ and a riskless reference point $r \in R$. The term $m\left(w\right)$ is the intrinsic "monetary utility" usually assumed relevant in economics, and the term $\mu\left(m\left(w\right) - m\left(r\right)\right)$ is the reference-dependent "gain-loss utility". This

[9] $f\left(\cdot\right)$ can be either continuous or discrete, or a mix of the two, over the domain $[0,1]$. In linear case, $f\left(\theta\right) = \theta$. The non-linear case can also be defined accordingly. For the most simplistic case, consider $f\left(\cdot\right)$ as a binary indicator, which equals to 1 if $\theta \geq \bar{\theta}$, and 0 otherwise. This suggests that a loan officer approves a loan when the perceived probability is greater than some cutting point $\bar{\theta}$, and rejects otherwise. Under different incentive schemes, a representative officer optimally chooses different $\bar{\theta}$ to guide his decision. This is also the setting used for the remaining analysis.

specification assumes that how a loan officer feels about gaining or losing relative to a reference point depends on the changes in monetary utility associated with such gains or losses (Koszegi and Rabin, 2006).[10]

When a loan officer is uncertain about loan outcomes, the reference point is a lottery over R, i.e. $F_d^t(\cdot)$ as discussed in the previous section:

$$U\left(w \mid F_d^t\right) = \int u\left(w \mid r\right) dF_d^t\left(r\right)$$

This formulation captures the notion that the evaluation of a wealth outcome is based on comparing it to all possibilities in the support of the reference lottery. A loan officer evaluates a stochastic wealth outcome w with some "mixed feelings" as the average of the different assessments $u\left(w \mid r\right)$ generated by $F_d^t(\cdot)$.

When the wealth outcome w is also drawn according to a lottery $F_{d'}^{t'}(\cdot)$, utility is given by:

$$U\left(F_{d'}^{t'} \mid F_d^t\right) = \int \int u\left(w \mid r\right) dF_d^t\left(r\right) dF_{d'}^{t'}\left(w\right)$$

This reference-dependent utility model (a) combines reference-dependent "gain-loss utility" with standard "monetary utility"; (b) bases the reference point to which outcomes are compared on endogenously determined beliefs and incorporates probabilistic beliefs; (c) allows for stochastic reference points.

A few caveats need to be mentioned here. First, up to this point, the model abstracts from non-linear probability weighting and assumes that preferences are linear in probabilities. The evaluation of a wealth outcome w is actually its expected reference-dependent utility. Second, despite of the assumption so far, the model of how utility depends on beliefs could be combined with any theory of how these beliefs are formed. In the next section, the model starts with the assumption that a loan officer correctly predicts the probabilistic environment and his own behavior in that environment, so that his beliefs fully reflect the true probability distribution of outcomes. The analysis then moves to how prior beliefs are formed and how they affect the decision behaviors. Third, the model specification can generate more realistic behavioral predictions. For example, a loan officer is less

[10]The separation and interdependence of economic and psychological payoffs is analogous to the assumptions made previously by Bell (1985), Loomes and Sugden (1986) and Kobberling and Wakker (2005). Discussion about functional forms and assumptions of monetary and gain-loss utility are available in the appendix.

averse to risk (of Type I error) with balanced adverse incentives and becomes more inclined approve a loan . Detailed discussion are available is the next two sections.[11]

4.4.3 Decision Strategies and Personal Equilibria

This section discusses a loan officer's decision strategies in equilibrium. Equilibrium cases here are defined differently by the time a decision is made. When a decision is made shortly before the outcomes occur, the concept "Unacclimating Personal Equilibrium" (UPE) defines the behavior where the stochastic outcome generated by utility-maximizing choices conditional on expectations coincides with expectations. The analysis further assumes a loan officer chooses his favorite UPE, the "Preferred Personal Equilibrium" (PPE). When a decision is made long before outcomes occur, a "Choice-Acclimating Personal Equilibrium" (CPE) defines a decision that maximizes expected utility given that it determines both the reference lottery and the outcome lottery. These definitions follow from Koszegi and Rabin (2006). Full description of the definitions are available in the appendix. Both PPE and CPE predict strong preference to insure the expected risk of penalty whenever possible. This prediction explains high rejection rates under the PRS and credit rationing problems in the financial market.

When a loan officer makes his lending decision shortly before the loan outcome is resolved, at that time the beliefs that determine the reference point are past and hence unchangeable. This means that he maximizes utility taking the reference point as given.

In this case, for a loan officer, choosing to approve a loan is optimal (i.e. a UPE) if $U\left(F_a^t|F_a^t\right) \geq U\left(F_r^t|F_a^t\right)$, where $t = 0, 1$ indicates the incentive schemes of pure PRS and PRS with additional penalty for Type II error respectively. More specifically, for example, under pure PRS, when $F_a = (\theta, B_a = 15, P_a = -20)$ and $F_r^0 = 0$, approving a loan is a UPE if:

$$[\theta 15 + (1-\theta)(-20)] + [\theta(1-\theta)\mu(35) + \theta(1-\theta)\mu(-35)] \geq [0] + [\theta\mu(-15) + (1-\theta)\mu(20)] \quad (6)$$

[11] Other previous literature, though closely related, has various limitations. For example, Sudgen (2003) compares outcome lotteries to reference lotteries state by state, capturing a form of state-contingent disappointment, but a state-independent feeling is missing. For another example, Bell (1985), Loomes and Sugden (1986), Gul (1991) and Shalev (2000) allow the existence of reference lottery but collapse it into some type of certainty equivalent. With this setting, reference lotteries that have some certainty equivalent generate the same risk preferences.

This shows a loan officer experiences direct "monetary utility" from the choice he made and a mixed-feeling of "gain-loss utility" based upon comparison between the choice he made and the reference lottery state by state. Both utility parts are represented by expected value. Given a reference lottery expectation, gain-loss feelings are generated whenever there is a gap between the outcome and the reference state. Given expectation to approve a loan, when an officer chooses to approve, he compares between a bonus of 15 and a penalty of -20; if instead he chooses to reject, he compares the riskless payoff 0 to the bonus and penalty cases.

Now, consider the case when a loan officer needs to make a loan decision and commit to it long before the loan outcome is resolved. In this case, the expectation relative to which an outcome of a decision is evaluated is formed in the future, and therefore incorporates implications of the decision.

Under pure PRS, approving a loan is optimal (i.e. a CPE) if $U\left(F_a^0|F_a^0\right) \geq U\left(F_r^0|F_r^0\right)$, i.e.:

$$[\theta 15 + (1-\theta)(-20)] + [\theta(1-\theta)\mu(35) + \theta(1-\theta)\mu(-35)] \geq [0] + [0] \tag{7}$$

In the case when there is additional penalty for Type II error, that is $F_r^1 = \left(\theta, P_r^1 = -20, B_r^1 = 0\right)$, approving a loan is a CPE if $U\left(F_a^1|F_a^1\right) \geq U\left(F_r^1|F_r^1\right)$, i.e.:

$$[\theta 15 - (1-\theta)20] + [\theta(1-\theta)\mu(35) + \theta(1-\theta)\mu(-35)] \geq [-20\theta] + [\theta(1-\theta)\mu(20) + \theta(1-\theta)\mu(-20)] \tag{8}$$

The difference between UPE and CPE is in the right-hand sides of inequalities 6 and 7, which capture the decision maker's expected utilities when deviating from the purported UPE and CPE, respectively. In UPE, the reference point does not adjust to the deviation, so rejecting a loan is assessed partly as a loss of \$15 (if the loan turns out to be good) and partly as a gain of \$20 (if the loan turns out to be bad). In CPE, the reference point does adjust to the deviation, so there is no sensation of gain or loss when rejecting a loan under pure PRS. This difference in feeling of gain and loss makes rejection a more favorable choice for a loan officer who needs to make lending decision long before the loan performance could be observed. Under the pure PRS situation, loan officers choose rejection to "insure" the potential risk of penalty.

4.4.4 Properties and Testing Hypotheses

In this section, three testing hypotheses of loan officers' lending behaviors are generated for empirical test.

$Hypothesis1:$ (Participation) For any incentive schemes F^t, where $t = 0,1$, if $U(w + F_a|w) \geq U(w|w)$, then $U\left(w + F_a|F^t\right) \geq U\left(w|F^t\right)$.

Hypothesis 1 has two implications. (a) If a loan officer is willing to approve a loan under a constant reference wealth level, he is also willing to make the same decision under some given incentive scheme based on risky loan outcomes. (b) The loan officer is at least not more risk-averse in the latter case than in the former.

Since loan officers, who are facing risky incentive schemes as background reference lotteries, are less risk-averse than those (normal people) who are not facing any risk [12], they are at least not less willing to make a lending decision than if they were in life. This guarantees positive loan supply in the financial market.

More broadly,

$Hypothesis1':$ If $U(w + F_a|w) \geq U(w|w)$, then $U\left(F^t + F_a|F^t\right) \geq U\left(F^t|F^t\right)$.

Intuitively, when a loan officer needs to make a decision for a new loan $d = a\,or\,r$, facing risky incentive schemes for all the loans he had made in the past makes him less averse to make another (risky) lending decision. When a new lending decision F_a is added to a riskless reference point w (as in the case of informal lending in life), positive outcomes of F_a are assessed as pure gains and negative outcomes of F_a are assessed as pure losses. But when F_a is added to the existing lottery under given incentive scheme F^t and evaluated relative to it (as the case a loan officer faces to) , positive outcome of F_a partially eliminate losses suffered from F^t in the past, and are hence evaluated more favorably than pure gains; similarly, negative outcomes of F_a partially eliminate gains from previous F^t, and hence, are evaluated less unfavorably than pure losses. For both this reasons, a loan officer who had made loans in the past under incentive scheme F^t is more willing to make another lending decision F_a. It can be concluded that expecting risk at the start decreases aversion to additional risk.

[12] Koszegi and Rabin (2007) interpreted this in a way that a person is less risk averse in eliminating a risk he expected to face than in taking on the same risk if he did not expect it.

Now consider the effect of a special incentive scheme - the PRS on decision behaviors. As discussed in section 2, PRS, as an on-going incentive scheme to control credit risk in most RCCs in China, gives a officer bonuses for the approved loans that are in performance and penalties for non-performing ones, but does nothing for the rejected loans. How does a loan officer behave under this setting? How does the behavior change if there are additional penalties for rejecting good loans?

$Hypothesis2:$ (Insure whenever Possible) If $U\left(w+F_a|w+F^0\right) \geq U\left(w|w+F^0\right)$, then $U\left(w+F_a|w+F^1\right) \geq U\left(w|w+F^1\right)$.

Hypothesis 2 tells us for any given loan application, if a loan officer is willing to approve under pure PRS (with reference lotteryF^0), then he will always be willing to approve it when there is additional penalties for Type II error of loan misclassification (with reference lottery F^1). With pure PRS, since there is no adjustment on wealth level for any rejected loans, a loan officer takes this as a "secure" option to avoid potential risk of penalty for bad loans, whereas with an additional dis-incentive of Type II error, this "secure" option is gone. As a result, a loan officer under pure PRS will never take higher risk than the case when there is additional penalties for Type II error. This "choosing insurance whenever possible" behavior under pure PRS makes loan officers be less willing to approve loans. Therefore, credit rationing exists in the financial market. Hypothesis 2 suggests that an additional dis-incentive for Type II Error can make a loan officer be more willing to approve a loan, and hence alleviates credit rationing in the market.[13]

In order to test the above hypotheses indicated by the decision behaviors, a field experiment is conducted with loan officers at local RCCs in China. Detailed information about experimental design is discussed in the next section.

[13]So far, all the results are based upon assumptions that first, people have linear probability weighting; and second, the probability belief θ is correct and fully reflects the distribution of risky outcomes. However, when these assumptions are loosen, a loan officer behaves differently. More specifically,

$Hypothesis3:$ (Ambiguity Increases Risk-Aversion) Loan officers are more conservative in making loan decisions when they are ambiguous about the application pool.

Ambiguity about the application pool makes positive outcomes less favorable and negative outcomes more painful, therefore, a loan officer is less likely to approve a loan given all others the same. Providing more information about loan applicants and guidance for prior perception formation may increase approval rates and alleviate credit rationing.

4.5 Experimental Design

4.5.1 Loan Officers

120 loan officers were invited to participate in the experiment via the internal staff systems of 3 local RCC Unions (RCCUs) at Heze City, Shandong Province in China.

Shandong is a coastal province which has relatively more farm and small businesses. As a result, it has larger population and higher representation of micro credit loans. Moreover, Shandong is also a typical province where the RCCs have been through effective structural reconstruction and institutional reforms in the past few decades. Heze City is located in the southwest of Shandong Province, and is a middle ranged city in terms of economy, population and geographic region. 3 out of 9 counties[14] in Heze City were selected to conduct the experiments, Chengwu, Cao and Shan County.

Each county has about 20 to 25 RCC branches (RCCs) that are open to the public and one union (RCCU) which provides supervision and administration. The 3 counties were each requested to invite 40 loan officers from their current staff to participate in the study, 10 officers per session, 2 sessions per day and 2 days in total. The three counties were also advised to list all their loan officers and assign invitations randomly. In the final selection of the loan officers, each branch had at least one loan officer participating in the study and some branches had 2 to 3. Demographics of the selected and non-selected officers were collected and compared to ensure representativeness.

4.5.2 Loan Files

10 loan files were selected from 2 branches of RCCs at Cao County and used as sample loans for evaluation in the experiment. The loans were previously approved within the past 5 years with repayment status known by the experimenter. The ratio of "good" versus "bad" loans were set at 50:50, that is 5 of the loans were performing loans (PLs) and the other 5 were non-performing loans (NPLs). In the experimental session, each participating loan officer was randomly given 6 out of the 10 loan files to evaluate. In order to make sure that the participating loan officers had never seen or known the loan files used in the study, the 2 branches which offered sample loan files were excluded when recruiting loan officers.

[14]Strictly, 8 counties and 1 district.

The 10 loan files were reformatted and reorganized by an identical structure. Each loan file was divided into 5 standardized parts. The first part included basic information about the applicant, such as application statement, age, education, employment, household background and business information, etc. The second part was a handwritten credit worthiness report prepared by the primary loan officer when the loan was first initiated. This report was prepared based on the loan officer's face-to-face interview with the loan applicant and on-site investigation of the business situation. The third part presented photocopies of all certificates, such as resident ID cards, marriage and employment certificates and documents of guarantors. The fourth part was standardized credit scores and credit reports printed from authorized credit institutes. And the last part was basic business information and credit reports for guarantors. Two guarantors were required for each loan application with amount below CNY 100,000 and one more guarantor for every CNY 50,000 increase in the loan amount.

Selection of 10 loan files balanced two considerations. First, it required enough number of loan files to generate variation for the loan file evaluation tasks. Second, it required enough repeats of evaluation for each loan file distributed across all treatments within the whole experiment, so as to make comparison and inference. Loan officers' evaluating different files and different parts differently in terms of behavioral responses, such as time use, confirmed the successful settlement of the first consideration. More details are available in discussion of robustness in Section 6.3. For the second consideration, choosing 6 out of 10 can generated 210 different combinations in total, but only 120 randomly drawn combinations were actually used for the study. By chance, no officer evaluated 6 loan files that were exactly the same as another. However, since each loan file on average appeared 72 times in total and 18 times within each treatment, these repeats enabled the tests of treatment effects. Actual frequencies of loan files in the experiment are provided in Section 5.1.

4.5.3 Treatments

A 2 by 2 between subject experiment was used. On one dimension, the treatment was whether or not there was additional dis-incentive (or penalty) for Type II error (i.e. rejecting good loans). On the other dimension, the treatment was whether or not the prior probability knowledge about the application pool was revealed to the loan officers.

Combining the two dimensions together yielded 4 treatment conditions. The *Baseline Condition*

(BC) had no penalty for Type II error (analogous to basic PRS) and no prior probability information about the application pool. It served as a benchmark for other treatments. The *Dis-Incentive Condition* (DC) had an additional dis-incentive for Type II error but still no prior probability information. The *Probability Condition* (PC) had basic PRS without penalty for Type II error but had the 50:50 prior probability information. The *Mixed Condition* (MC) had both additional dis-incentive and prior probability information.

To be more specific, on the dis-incentive dimension, when there was no penalty/dis-incentive for Type II error, officers were offered the incentive structure $(15, -20, 0, 0)$. It meant:

- getting CNY 15 as bonus for every approved performing loan (PL),
- losing CNY 20 as penalty for every approved non-performing loan (NPL) (i.e. Type I error),
- getting nothing for the rejected loans, regardless of their performance status.

In contrast, when there was dis-incentive for Type II error, the incentive structure changed to $(15, -20, -20, 0)$. It meant:

- getting CNY 15 as bonus for every approved performing loan (PL),
- losing CNY 20 as penalty for every approved non-performing loan (NPL) (i.e. Type I error),
- losing CNY 20 as penalty for every rejected PL (i.e. Type II error),
- getting nothing for rejected NPL.

Comparing the two conditions, no penalty for Type II error represented the basic Personal Responsibility System (PRS) which controlled financial credit risk by penalizing Type I error only. An additional dis-incentive (penalty) for Type II error was represented by the third (non-zero) number in the incentive structure, i.e. -20 in this case. It would be interesting to show under the PRS, risk-averse officers tended to avoid potential penalty by rejecting any suspicious loans and hence, resulted in credit rationing problem. However, with an additional dis-incentive for Type II error, loan officers no longer had the "secure" option to avoid penalty and hence, put more cognitive effort in loan classification and became more likely to approve loans. Credit rationing problem could then be alleviated.

One thing that needs to be clear is these payoff structures were imposed upon CNY 150 participation fee. By the design of the study, each loan officer could get as low as CNY 30 or as high as CNY 270 depending on their evaluation accuracy. The expected payoff was CNY 165, which was roughly an loan officer's daily salary.

On the probability dimension, when it was not treated, the proportion of PLs and NPLs were not revealed to the loan officers. When it was treated however, the officers were told that the six files that they were going to evaluate were randomly selected from a loan database with 50:50 PLs and NPLs. With the treatment of this prior probability information, results in the study can shed light on how ambiguity and prior beliefs affected loan decisions.

4.5.4 Experimental Procedures

At the beginning of the experiment, treatment condition was first randomly assigned to each session. Each county had a full circulation of all 4 treatment conditions with random order. Within each session, the participating loan officers were given identical instructions before they performed any tasks. The instruction consisted of 3 paragraphs, with the first one addressing the resource of the loan files and definitions of PLs and NPLs, the second one explaining the payoff structure and the third one outlining the standardized loan file structure. Instructions for the 4 treatment conditions were otherwise identical except for the prior probability indication in the first paragraph and the incentive structure in the second paragraph.

10 identical lap top computers were purchased to conduct the experiment. A visual basic program based on excel was designed and installed on each computer to guide through 6 loan evaluation procedures from the beginning to the end. The 10 standardized loan files were also saved in the computers and linked to the program. At the beginning of each session and before the loan officers arrived, the experimenter initiated the program on each computer by inputting the treatment and the loan file numbers that were randomly generated ahead of time. After all 10 officers arrived, the experimenter explained the procedure and instructions to all and used an eleventh sample loan file to show how to do the evaluation on a computer. The loan officers were advised not to operate computers until they were told to, but was allowed to ask questions if they found anything unclear. When all officers were clear about the procedure and had no more questions, the experiment started.

The first page a loan officer saw on the computer was an instruction with the 3 paragraphs

mentioned above. The second page listed all the warnings, such as no discussion, no tempting to quit the program, and indicating the experimenter if need to stop, etc. A loan officer was then directed to a loan evaluation page. The loan evaluation page appeared 6 times in a row, each representing one loan file that was randomly selected for evaluation.

On the loan file evaluation page, there were 5 buttons that can navigate to the corresponding part of a loan file. A loan officer needed to press the buttons and review the 5 parts in order. After reviewing each part, a loan officer was asked to provide a score (an integer number between 1 to 9) indicating his tendency to approve the loan so far, where 1 representing definitely inclined to reject and 9 definitely inclined to approve. It was only when a loan officer finished reviewing one part of a loan file and reported his approval tendency, can he continue to the next part. After reviewing all 5 parts and providing tendency scores, a loan officer was given a chance to review the full application with all 5 (separate) documents attached together before they made their final decisions (approve or reject). The loan officers were told that the tendency scores were for reference only and would not affect their final payoffs, but their final decisions would. Neither the tendency scores nor the final decisions could be changed once they had been submitted.

Another uniqueness of the design was that the program recorded the time when a loan officer pressed a button to review a standardized document and the time when he clicked a button to submit his score or decision after reading the document. The time spent on each document can then be derived by calculating the difference between a pair of the time records. Absolute time used revealed the efforts put in the evaluation tasks. The percentage time used (out of total time spent on a loan file) revealed the perceived relative importance of the corresponding document. An optional 15 mintues' break was built in the program after a loan officer finished reviewing 3 loan files.

After all 6 loan file evaluations, a loan officer was asked to finish a computer-based survey to complement the study. The survey included questions for demographics, knowledge about the Personal Responsibility System (PRS), perceptions about lending accuracy, self confidence and risk aversion, etc. Screen shots of the computer program, front page of a sample loan file and the electronic survey are provided in the appendix.

4.5.5 Identification Strategy

Average treatment effects (ATEs) can be identified by the mean differences between treatment conditions. Specifically, behavioral responses can be investigated by the following model specification:

$$Y_{ij} = \theta_1 \cdot Dis-Inct_{ij} + \theta_2 \cdot Prob_{ij} + t_0 \cdot TimeTot_{ij} + \sum_{k=1}^{5} t_k \cdot TimeRel_{ij}^{k} + \sum_{k=2}^{6} \delta_k \cdot Order_{ij}^{k} + \alpha_i + \beta_j + u_{ij}$$

- Y_{ij} represents behavioral responses (i.e. lending decision, occurrence of Type I error, etc.) for loan file i in the jth evaluation (during the whole experiment), where $i = 1, ..., 10$ and $j = 1, 2, ...$
- $Dis - Inct_{ij}$ and $Prob_{ij}$ are dummy variables indicating treatment status of dis-incentive for Type II error and prior probability information respectively.
- $TimeTot_{ij}$ is total time used to evaluate a loan file.
- $TimeRel_{ij}^{k}$ is relative proportion of time used in reviewing part k, where $k = 1, 2, 3, 4, 5$. Note that the 6th part, i.e. the final review was omitted.
- $Order_{ij}^{k}$ is a group of dummy variables indicating whether or not the $ijth$ loan evaluation happened as the kth file being evaluated within a session, where $k = 2, 3, 4, 5, 6$. This group of dummies control for order effects. Loan file being evaluated as the first one in a session was set as default.
- α_i is the loan file fixed effect.
- β_j is the fixed effect for loan file repeats across all 12 sessions .

With this setting, estimates of θ_1 and θ_2 are the ATEs of *Dis-Incentive Condition* (DC) and *Probability Condition* (PC) respectively, with the sum of the two being the ATE of the *Mixed Condition* (MC).

4.6 Results

The empirical analysis in this section proceeds in 3 steps. After a brief introduction of the dataset used in this part, results first show how incentive mechanism and probability perception affect lending decisions. Credit rationing behavior, profitability of the financial institutes and loan supply in the market are discussed accordingly. In the second step, results explore the effects of behavioral responses, such as efforts, confidence and risk aversion, on loan decision and how they are related to the exogenous treatments. Lastly, credit ratings for loan files are compared across treatment conditions. Inferences about lending decisions and accuracy are made from there.

4.6.1 Summary Statistics

The dataset used for analysis came from 3 major sources. The primary part of the data was collected during the experiment, which contained credit ratings of 5 standardized loan file parts, time used in the corresponding parts and the final lending decisions for 720 loan file evaluations. The dataset was then matched with the survey finished by the 120 loan officers after the experiment. The survey included information such as loan officers' demographics, knowledge about the Personal Responsibility System (PRS), perceptions about lending accuracy, self confidence and risk aversion, etc. Finally, business information about the 10 sample loan files used in the experiment was recovered and complemented the dataset.

Table 21 lists demographics of the participating loan officers. 76 out of 120 loan officers are male, which accounts for 63.3% of the sample. The age ranges from 22 to 48, with average age at 34. The average educational level is 2-year college. Due to social development in the past few decades, some younger officers under 30 have degrees as high as graduate level, while some older officers above 40 maintain an educational level as low as middle school. The mean income level is around CNY 3,000 to 4,000 per month. Participants' professional training and working experience are widely dispersed, ranging from one or two months to 30 years. On average, they have about 2.5 years' professional training and roughly 3 years' working experience as loan officers. These statistics match the loan officer population group of RCCs and even other financial institutes in China. Furthermore, decomposing the summary statistics by treatment, there is no significant difference between treatments. As a result, both representativeness and randomization are obtained.

Table 21: Study 3_Summary Statistics of Demographics

Table: Summary Statistics for Loan Officer Demographics

	No Penalty for TypeII		Penalty for TypeII Error	
Variable	Mean	SD	Mean	SD
Male	0.633	0.486	0.633	0.486
Age	34.600	7.441	34.283	7.531
Edu	3.117	0.739	3.133	0.566
Training	2.518	4.652	2.629	4.920
Experience	5.161	6.834	4.380	5.200
Income	1.667	0.601	1.667	0.795
Obs.	60		60	
	Not Know Probability		Know Probability	
Variable	Mean	SD	Mean	SD
Male	0.600	0.494	0.667	0.475
Age	34.617	7.434	34.267	7.537
Edu	3.117	0.524	3.133	0.769
Training	1.723	3.352	3.425	5.757
Experience	5.190	7.010	4.351	4.956
Income	1.683	0.651	1.650	0.755
Obs.	60		60	

The 10 sample loan files used in the experiment are all farm or small business loans with amount ranging from CNY 30,000 to 200,000. The duration of loan is between 12 to 24 months. The applicants on average come from household size of 3 to 4, with age between 25 to 57 at the time when they applied for loans. The average annual net income per capita is about CNY 20,000. On the days of experiment, each of the 120 loan officers was randomly assigned 6 out of the 10 sample loan files to evaluate. On average, each loan file should exist 72 times, 18 times in each of the 4 treatment conditions. Table 22 lists the actual evaluation frequency of all 10 files from random generating processes. Loan files numbered from 1 to 5 are performing loans (PLs) and those numbered from 6 to 10 are non-performing loans (NPLs). Each loan file appeared between 57 to 81 times in total and between 12 to 24 times in one section.

The mean payoffs are shown in Table 23. On average, a loan officer got CNY 163 for participating in the study. Comparing to average monthly income of CNY 3,000 to 4,000, this final payoff is roughly equal to daily salary. Payoffs are about CNY 11 lower in treatments with additional dis-incentives for Type II error (*Dis-Incentive*, DC and *Mixed Condition*, MC) than treatments with pure PRS (*Baseline*, BC and *Probability Condition*, PC). The former is CNY 158 and the latter is CNY 169. However, the difference is not statistically significant. Since the payoff structure

Table 22: Study 3_Loan File Evaluation Frequency

Table: Loan File Evaluation Frequency

Loan Nature	fileNO	Baseline	Dis-Incentive	Probability	Mixed	All	
	1	24	21	21	20	86	
	2	13	14	14	17	58	
Performing	3	22	17	21	20	81	
Loans	4	14	20	13	12	59	
	5	17	18	21	21	77	361
	6	14	23	20	19	76	
	7	20	14	12	15	61	
Non-Performing	8	21	17	20	21	78	
Loans	9	17	15	16	15	63	
	10	18	21	22	20	81	359

Table 23: Study 3_Final Payoffs by Treatment

Table: Loan Officers' Final Payoffs by Treatment (in CNY)

	Variable	Obs	Mean	Std. Dev.	Min	Max
ALL	payment	720	163.54	21.51	100	195
Baseline	payment	180	169.00	14.50	130	195
Dis-Incentive	payment	180	159.50	24.37	100	195
Probability	payment	180	168.50	14.77	135	195
Mixed	payment	180	157.17	26.90	100	195

performed as performance salary for loan officers, the insignificance of payoffs between treatments implies that decision accuracy is higher in DC (and MC) so that it can cover up the additional dis-incentives. Probability treatment does not yield difference in payoffs.

4.6.2 Credit Rationing, Loan Decision and Accuracy

Out of the 720 loan evaluations, 375 were approved and 345 were rejected. The approval rate is 52%. 75% of the decisions were correct, which means 578 loans were correctly identified and 178 were misclassified. Further, out of the 375 approved loans, 279 were PLs and 96 were NPLs. The probability of Type I error (i.e.approving NPLs) is 25.6% conditional on approval. Correctly approving PLs implies generating interest revenues, however, Type I error implies losing principals. In contrary, out of the 345 rejected loans, 278 were correctly rejected (as NPLs) and 82 PLs were misclassified, yielding Type II error (i.e. rejecting PLs). The probability of Type II error is 23.77% conditional on rejection. While correctly rejecting NPLs avoids losing principals, Type II error makes the financial institutes foregone potential revenues. Table 24 provides a full description of these rates.

Table 24: Study 3_Frequency of Decision Accuracy

Table: Frequency of Decision and Accuracy

		Baseline	Dis-Incentive	Probability	Mixed	All	%
Decision	Approve	80	91	93	111	375	0.52
	Reject	100	89	87	69	345	0.48
Accuracy	Correct	133	141	135	133	542	0.75
	Wrong	47	39	45	47	178	0.25
Category	Correct_Good	62	71	69	77	279	0.39
	Correct_Bad	71	70	66	56	263	0.37
	TypeI	18	20	24	34	96	0.13
	TypeII	29	19	21	13	82	0.11
Total Obs.		180	180	180	180	720	1.00

In addition, Table 24 lists the frequency distributions across 4 treatment conditions. Number of approval is about 10 (loans) higher in the *Dis-Incentive* (DC) and the *Probability Condition* (PC) than in *Baseline Condition* (BC). In the *Mixed Conditions* (MC), the gap is enlarged to 31 (loans). Number of correct decisions is 8 (loans) higher in DC than other conditions. Type II error is much lower in DC, PC and MC than BC, however, Type I error is higher in MC than the other three conditions.

Table 25 lists the panel logistic regression results of average (marginal) treatment effects (ATE) on 6 lending decision measures. Each column of Table 25 represents regression of one decision measure. The 6 decision measures are binary indicators of

1. Whether a loan was approved or not, $Approve = 1$ if yes, 0 otherwise.
2. Whether a loan decision was correct or not, $Correct = 1$ if yes, 0 otherwise.
3. Whether a loan decision correctly approved a performing loan, $Correct - Good = 1$ if yes, 0 otherwise.
4. Whether a loan decision mistakenly rejected a performing loan, $TypeII = 1$ if yes, 0 otherwise.
5. Whether a loan decision correctly rejected a non-performing loan, $Correct - Bad = 1$ if yes, 0 otherwise.
6. Whether a loan decision mistakenly approved a non-performing loan, $TypeI = 1$ if yes, 0 otherwise.

Treatment dummy $Penalty - TypeII$ indicates whether the incentive structure had penalty for

Table 25: Study 3_ATE on Decision and Accuracy

Table: Average Marginal Treatment Effects on Decision and Accuracy by Panel Logistic

	1	2	3	4	5	6
	Approve	Correct	Correct_Good	TypeII	Correct_Bad	TypeI
Penalty TypeII Error	0.129***	0.016	0.102**	-0.102**	-0.081**	0.081**
	(0.045)	(0.032)	(0.044)	(0.044)	(0.047)	(0.047)
Know Probability	0.120**	-0.014	0.077*	-0.077*	-0.087*	0.087*
	(0.046)	(0.033)	(0.045)	(0.045)	(0.047)	(0.047)
Control for						
Loan File Fixed Effect	yes	yes	yes	yes	yes	yes
Order Effect	yes	yes	yes	yes	yes	yes
Total Time Used	yes	yes	yes	yes	yes	yes
Relative % Time Used	yes	yes	yes	yes	yes	yes
Log Likelihood	-396.06	-378.52	-181.28	-181.28	-182.73	-182.73
LR Chi_2(DF)	258.76(22)	48.25(22)	54.29(17)	54.29(17)	51.45(17)	51.45(17)
Observations	720	720	361	361	359	359

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Type II error or not, 1 if yes and 0 otherwise; similarly, $Know-Probability$ indicates whether prior probability information about loan application pool was revealed or not, 1 if yes and 0 otherwise. Entries of the first two rows in Table 5 are estimated marginal effects of the treatment dummies on corresponding decision measures. The combination of controlled variables used in each regression is listed below the estimated marginal effects.

Column 1 of Table 25 is the regression result about approval. The ATEs on approval reveals credit rationing between treatments. By definition, credit rationing happens if given loan applicants that appear equal, some receive a loan and some do not (even when they offer to pay a higher interest rate) (Stiglitz and Weiss, 1981). When controlling for loan file fixed effects, the ATEs on probability to approve a loan represents credit rationing. Compared to the *Baseline Condition* (BC), in the *Dis-Incentive Condition* (DC), the approval rate (or the probability to approve a loan) is 13% higher. This result suggests credit is rationed by 13% on average under the Personal Responsibility System (PRS). Further, in the *Probability Condition* (PC), the approval rate is 12% higher than BC. Providing prior probability information to loan officers alleviates credit rationing by 12%. When combining the two treatments, in the *Mixed Condition* (MC), the rate on average is 25% higher than that in BC. When a loan officer faces additional dis-incentive (or penalty) for Type II error and at the same time knows the prior probability about loan application pool, he becomes 25%

higher in probability to approve a same loan than he would otherwise.[15]

Column 2-6 are regressions about loan decision accuracy. It shows that Type II error is 9.5% lower in DC and 6.1% lower in PC than in BC, given a loan being a PL. These probabilities also implies the percentage of credit being rationed from profitable clients (customers). A decrease in Type II error implies a decrease of foregone revenues, and hence an increase of interest returns. In contrast, given a loan being a NPL, Type I error is 8.2% higher in DC and 4.4% higher in PC. An increase in Type I error implies an increase in loss of principals.[16]

4.6.3 Time Use, Confidence and Risk Aversion

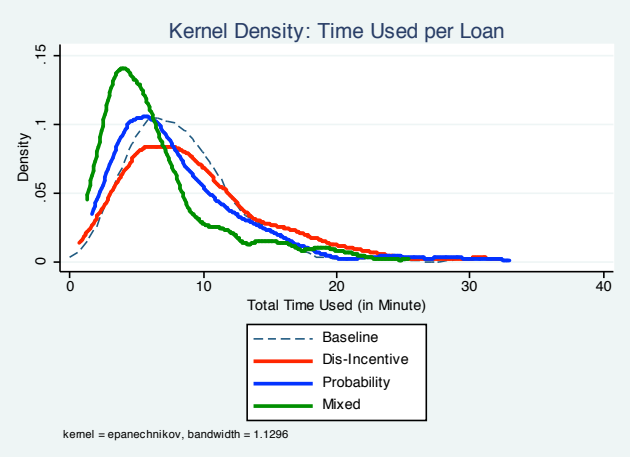
Time used in reviewing each part of a loan file was recorded by the computer program during the experiment. Total time used for one loan file evaluation was then calculated by summing up the time used for all 5 separate parts and the time used for the final full document review before a loan decision was made. On average, it took about 8.3 minutes (8 minutes and 18 seconds) for a loan officer to evaluate a loan file. The longest time recorded for a single loan evaluation was 32.93 minutes.

Figure 5 sketches the estimated kernel density distributions of total time used for the 4 treatment conditions. All curves are skewed to the right. The median for the *Baseline Condition* (BC) is 7.83 minutes. Dis-incentive treatment increases the median to 8.13 minutes and the probability treatment decreases the median to 6.85 minutes. Combining the dis-incentive and the probability treatment further decreases the time use to 5.26 minutes. These changes reveals officers' perceptions and attitudes towards loan files. Comparing to BC, additional dis-incentive makes officers devote

[15]Total credits approved by each treatment condition are calculated as market level proxies. In BC, CNY 3,950,000 was granted. In DC, it was CNY 5,050,000. In PC and MC, the number was CNY 5,100,000 and CNY6,340,000 respectively. If the treatment conditions are analogous to corresponding market situations, the market level credit rationing under the PRS could be as high as 37.7%. This implies that almost 40% of the funding that should have been available to the market is rationed due to the current incentive mechanism.

[16]However, one should not draw a conclusion of the overall effects by merely comparing the changes in Type I and Type II error. Instead, profits realized under each treatment conditions should be investigated. Based on the loan specific information and officers' lending decisions, net gains in the 4 conditions are derived as firm level profits of RCCs. As shown in Appendix Table 1, BC has the lowest revenue, the highest cost for Type II error and second highest cost for Type I error out of the 4 conditions, which yields up to roughly 41.52% decrease in profit (as oppose to the most ideal case). Note that numbers used are net present values. The interest rate is identical for all loans, which is 1.1695% per month. Revenue from approving PLs is the interest returns. Cost of Type I error is set to the loan principal. Cost of Type II error is calculated as a recursive function of cost of Type I as proposed by Nayak and Turvey (1997), i.e. $Cost_{TypeII} = r - (\theta r - (1-\theta)\, Cost_{TypeI})$, where r stands for interest return, θ stands for the expected probability of repayment. Finally, the revenues and the cost are rescaled according to their corresponding probabilities for comparison purpose.

Figure 5: Study 3_Kernel Density of Time Used per Loan



more cognitive efforts (and hence, more time) so as to avoid any loan misclassification. Probability treatment removes ambiguity about the application pool. As a result, it takes less time for a loan officer to make loan decisions. In the *Mixed Condition* (MC), under the combination of prior probability information and additional dis-incentives, concerns of NPLs are no longer as important as in other conditions, as a result, the time use is further decreased.

Figure 6 describes how time use changes during one experimental session of 6 loan file evaluations. In general, for all the 4 treatment conditions, the average time use declines as officers evaluate more loan files, from 10-15 minutes for the first file to less than 10 minutes in the last file. In addition, the changing pattern suggests a sharper decline at the beginning for the first 1 or 2 files, and a more steady trend for the remaining files. Comparing the 4 treatment conditions, the sharpe decline of time use happens at the third file in BC, DC and MC, but occurs as early as in the second file in PC. This difference further supports the implication that probability treatment removes ambiguity and makes the officers spend less time on loan applications.

Figure 7 lists the percentage of time spent on different parts of a loan file. These percentages reveal the relative importance perceived by loan officers. On average, loan officers spend roughly 25% of the time reviewing the first part, which includes the basic business information and household economic background of the applicants. Another 20% of the time is spent on guarantors' information (part 5), including guarantors' business information, household situation and their standardized credit scores and reports. The third important part is the handwritten credit worthiness investigation report made by the primary loan officer at the time when the application was initialized.

Figure 6: Study 3_Total Time Used per Loan across Rounds

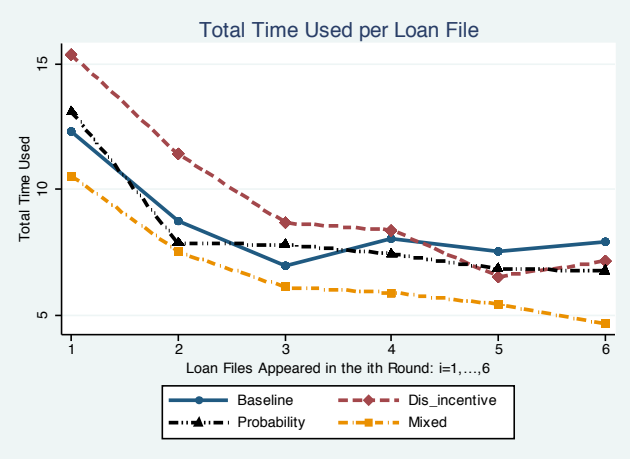
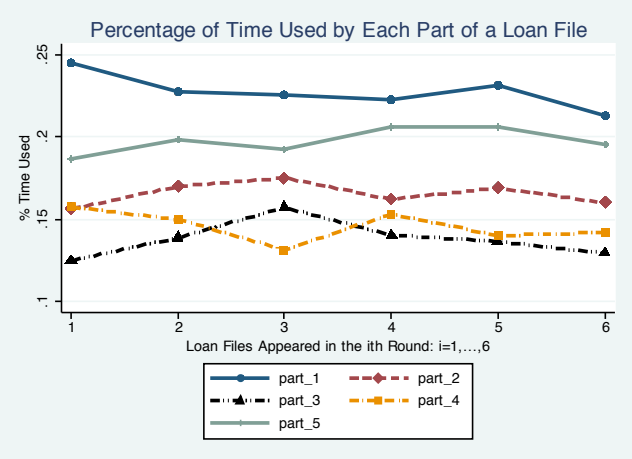


Figure 7: Study 3_Percentage of Time Used by Part



This part takes about 16% of the total time (part 2). It is then followed by the applicant's and spouse's credit scores and reports, about 14%. The least important part perceived by the officers is ID certificates, i.e. marriage and employment, etc, which takes about 13% of the total time. Note that the remaining percentage is for the full document review before final lending decision was made. Comparing the proportions across 6 files during one experimental session, the percentages are stable, which implies that the perceived relative importance is consistence within all the 6 loan file evaluations.

In addition to time efforts, loan officers' confident level and risk-aversion are estimated through groups of psychometric questions in the survey. Questions that are used to elicit confidence levels are, for instance, "how do you think your loan portfolio relative to other colleagues?" and "how

confident you feel when you evaluate loan files at work?", etc. Questions that are used to reveal risk aversion are, for example, "how much are you concerned about approving NPLs and getting penalized?" and "how much are you concerned about losing you job if in case you approved NPLs?", etc. The response of questions ranges from 1 to 5, with 1 suggesting the least and 5 the most. Average values of the group questions are then calculated as proxies for confidence and risk-aversion. [17]

Table 26 provides the marginal effects of time use, confidence and risk aversion on 6 decision measures. Each column represents one regression. The decision measures follow the definition in Table 5. Explanatory variables are total time used in evaluating a loan file, percentage time used in each of the 5 parts, indexes of confidence and risk aversion. Combination of control variables are listed in the bottom part. Entries of Table 26 are estimated marginal effects derived from panel logistic regressions.

Based on Table 26, loan decision is significantly affected by total time use, relative time use in part 2 (credit worthiness report prepared) and risk aversion. For every 1 more minute used in reviewing a loan file, the probability to approve decreased by 1.65%. In this sense, the total time use suggests some level of hesitation. When a loan officer spent 1 more percent of total time on part 2, the probability to approve this loan increases 8.62% . Risk aversion imposes a negative effect on approval. For decision accuracy, confidence level negatively impacts the probability of a correct decision. To some extend, this confidence measure represents some over-confidence level. Furthermore, longer total time use and higher risk aversion imply higher Type II error. Higher confidence yields higher Type I error. Spending more time on applicants' credit report (Part 4 of a loan file) decreases the Type I error.

4.6.4 Credit Rating and Tendency to Approve

During the loan evaluation process, loan officers were asked to provide a score after reading each part of a loan file. The score indicates the tendency to approve a loan based on the information a loan officer had read so far, with 1 representing definitely inclined to reject and 9 definitely inclined to approve. In total, there are 5 scores recorded for each loan file. Note that as a loan officer reviews more and more information, the score indicates his "cumulative" tendency to approve a loan. The

[17] Some may argue that time use may also be impacted by confidence and risk aversion. Actually, this is not true. Scatter plot with linear fit of time use against confidence level (or risk aversion) shows there is no significant relationship between the two. Hence, time use suggests effort only. Detailed results are available upon request.

Table 26: Study 3_Marginal Effects of Time Use on Decision

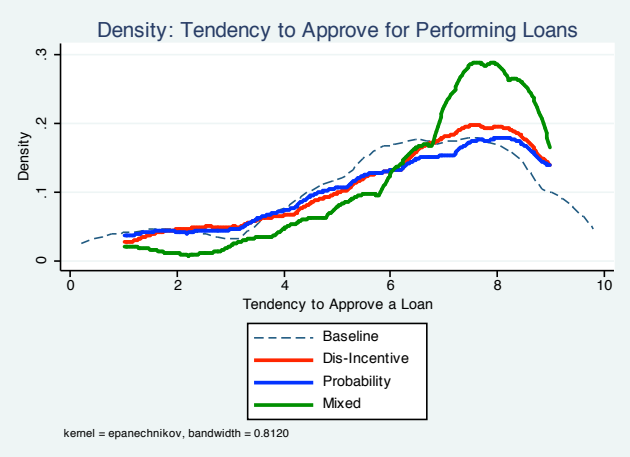
Table: Marginal Effects of Time Use, Confidence and Risk Aversion on Decision by Panel Logistic

	1 Approve	2 Correct	3 Correct_Good	4 TypeII	5 Correct_Bad	6 TypeI
Total Time Used (in Min)	-0.0165***	-0.0014	-0.011**	0.011**	0.0067	-0.0067
	(0.006)	(0.004)	(0.005)	(0.005)	(0.007)	(0.007)
% Time Used in Part 1	-0.403	0.009	-0.28	0.28	0.323	-1.323
	(0.297)	(0.202)	(0.309)	(0.309)	(0.263)	(0.263)
% Time Used in Part 2	-0.862**	0.356	-0.286	0.286	0.581	-0.581
	(0.417)	(0.283)	(0.421)	(0.421)	(0.396)	(0.396)
% Time Used in Part 3	-0.156	0.169	0.096	-0.096	0.296	-0.296
	(0.395)	(0.276)	(0.389)	(0.389)	(0.382)	(0.382)
% Time Used in Part 4	-0.512	0.249	-0.146	0.146	0.637*	-0.637*
	(0.331)	(0.231)	(0.313)	(0.313)	(0.330)	(0.330)
% Time Used in Part 5	-0.111	0.265	0.104	-0.104	0.281	-0.281
	(0.302)	(0.207)	(0.281)	(0.281)	(0.300)	(0.300)
Confidence	0.061	-0.053*	-0.0101	0.0101	-0.083*	0.083*
	(0.044)	(0.031)	(0.040)	(0.040)	(0.046)	(0.046)
Risk-Aversion	-0.083**	-0.0074	-0.061*	0.061*	0.048	-0.048
	(0.034)	(0.024)	(0.033)	(0.033)	(0.034)	(0.034)
control for						
ATE	yes	yes	yes	yes	yes	yes
Loan File Fixed Effect	yes	yes	yes	yes	yes	yes
Order Effect	yes	yes	yes	yes	yes	yes
Loan Officer Demo.	yes	yes	yes	yes	yes	yes
Log Likelihood	-357	-372.28	-175.74	-175.74	-170.32	-170.32
LR Chi_2(DF)	282.87(34)	60.87(34)	75.37(29)	75.37(29)	76.27(29)	76.27(29)
Observations	720	720	361	361	359	359

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Figure 8: Study 3_Tendency to Approve - Performing Loans

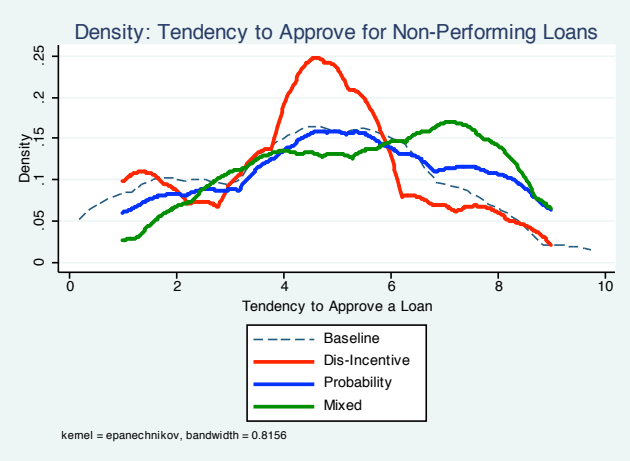


final score (score after reading all 5 parts) fully reflects the loan decision.

Figure 8 and 9 provide the kernel densities of the final score for performing loans (PLs) and non-performing loans (NPLs) respectively by 4 treatments. On average, officers rated 6.43 (out of 9) for PLs and 4.87 for NPLs. The significant difference between the ratings (F=90.09, P=0.00) suggests loan officers can differentiate between PLs and NPLs very well, which confirms validity of the selected loan files. More detailed discussion is provided in Section 16.2. For both PLs and NPLs, the *Probability Condition* (PC) and the *Mixed Condition* (MC) have larger probability density for higher scores than the *Baseline Condition* (BC). This yields higher mean or median of PC and MC than BC, with the former about 0.42 points higher (F=2.78, P=0.096) and the latter about 0.94 points higher (F=15.82, P=0.00). This is another evidence which supports Hypothesis 3 that prior probability information removes ambiguity and makes officers more inclined to approve loans.

By Hypothesis 2, additional dis-incentive treatment would made officers more serious (or conservative) in rejecting loans so as to avoid penalty for Type II error. If this were true, one would find officers became less willing (or possible) to rate low scores and/or more willing (or possible) to rate higher scores in DC than in BC. These implications are supported by Figure 8 and 9. In Figure 8, for PLs, the curve of DC slightly first order stochastic dominates that of BC, i.e. less mass for lower scores and more mass for higher ones in DC than BC. This difference suggests higher tendency to approve (performing) loans. In Figure 9, for NPLs, some mass for lower rates in BC now moves towards middle in DC, however, due to the non-performing nature of these loans, the mass for higher scores does not increase as it does for PLs. As a result, the curve becomes a tighter

Figure 9: Study 3_Tendency to Approve - Non-Performing Loans



distribution with relatively lower variance in DC than in BC for NPLs. This suggests loan officers are more conservative (and hence, less possible) to reject loans in the *Dis-Incentive Condition* than the baseline due to the additional penalty for Type II errors.

Table 27 lists regression results of 5 reported approval tendencies and the final decision. Each column represents regression of one tendency. Explanatory variables include the two treatment dummies, absolute cumulative time used by the time when a tendency was reported and relative time used (in percentage) on loan file parts that had been reviewed so far. Combination of control variables is listed on the bottom. According to Table 27, due to the removal of ambiguity about the application pool, probability treatment results in significant higher tendency to approve immediately from the first part until the last. Dis-Incentive treatment only performs significant positive impact on the final score and the loan decision. It is only when it comes to the final decision, does the penalty for Type II error become a concern to the loan officers. Effects of time use on score (or tendency to approve) are universally negative on all tendency reports, consistent to the result in Section 15.3 about (discrete) lending decision. Furthermore, this negative effect holds for both self-effects from current parts (diagonal coefficients) and cross-effects from all previous parts (upper diagonal coefficients).

4.7 Robustness

This section provides discussion about the experimental validity. Potential concerns come from 3 sources: first, the randomness and representativeness of participants; second, the successful selection

Table 27: Study 3_ATE on Tendency to Approve a Loan

Table: Average Treatment Effect (ATE) on Tendency to Approve a Loan

	1	2	3	4	5	6
	Tendency to Approve a Loan after (Reading)					
	Part 1	Part 2	Part 3	Part 4	Part 5	Final Review
Penalty for TypeII	0.195	0.122	0.127	0.137	0.330**	0.0920***
	(0.157)	(0.155)	(0.158)	(0.159)	(0.163)	(0.032)
Know Probability	0.503***	0.480***	0.670***	0.301*	0.496***	0.0744**
	(0.161)	(0.159)	(0.162)	(0.162)	(0.167)	(0.033)
Time Used after						
Part 1	-0.101					
	(0.088)					
Part 2		-0.0462				
		(0.055)				
Part 3			-0.149***			
			(0.043)			
Part 4				-0.117***		
				(0.032)		
Part 5					-0.0967***	
					(0.025)	
Final Review						-0.0108**
						(0.004)
% Time Used in						
Part 1	-2.223**	-1.104	0.0625	-0.918	-1.29	-0.22
	(0.978)	(0.894)	(0.894)	(0.899)	(0.953)	(0.199)
Part 2		-1.146	-0.363	-1.362	-0.478	-0.484*
		(1.231)	(1.242)	(1.244)	(1.288)	(0.260)
Part 3			-3.549***	-0.821	-1.389	-0.11
			(1.235)	(1.237)	(1.283)	(0.258)
Part 4				-2.812**	-1.032	-0.36
				(1.142)	(1.169)	(0.232)
Part 5					-0.141	-0.0512
					(1.054)	(0.207)
Control for						
Order Effect	yes	yes	yes	yes	yes	yes
Loan File FE	yes	yes	yes	yes	yes	yes
Loan Officer FE	yes	yes	yes	yes	yes	yes
Observations	720	720	720	720	720	720
R-squared	0.204	0.261	0.38	0.266	0.223	0.339

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

of sample loan files; and third, learning or order effects during the whole study period.

4.7.1 Experimental Validity: Demographics

As discussed in Section 15.1, there was no significant difference in demographics of loan officers across treatments. Randomization of treatment assignments among participants is guaranteed. Further, in order to absorb all potential effects of demographic variables on the results, estimations are made with these demographics as control variables. Significant impacts only come from educational level and working experience for Type I error. The former has a negative impact and the latter positive[18]. Moreover, the results of interest in this study are consistent when estimated with or without demographic variables. Detailed discussion of consistency are available in Section 16.4. In order to ensure representativeness, statistics for demographics are compared among the groups of experiment participants, employees of RCC's in Heze City and employees in the financial system in China. No significant difference is found and representativeness is established.

4.7.2 Experimental Validity: Loan File Fixed Effects

The successful selection of sample loan files should have two major features. First, any loan file should not be identified to be a performing or non-performing loan for sure. More strictly, no loan file should be considered as good or bad with significantly high probability. This feature guarantees the validity of causal effects between treatments and loan decisions. Second, loan files should be randomly selected from the population pool so that they provide enough variation for participating loan officers to response.

Test of the first feature can be conducted by summarizing decision and accuracy measures by loan file and checking if the mean rates for any individual loan file is greater than 90% or smaller than 10%. Appendix Table 2 lists the percentage of correct classification for each loan file. The percentages range from 55.13% to 86.84%. No loan file is considered with very high probability to be PL or NPL.

Appendix Table 3 provides the estimated coefficients of regressing the 6 decision measures on loan file dummies, with a performing loan, File No.1, as the default case. Each column of the table represents regression of one decision measure. The definitions of decision measures follow from Table

[18]Regression results omitted due to space limit.

25 in Section 15.2. For the lending decision, i.e. approve or not (Column 1), files are only different between PLs (File No.1 to 5) and NPLs (File No.6 to 10), with the former group being approved with higher probabilities than the latter. Consistent with the approval tendency score in Section 5.4, results here suggest that loan officers can differentiate between PLs and NPLs. The loan files are not biased. Furthermore, loan files do not make any difference in terms of correctness (Column 2 to 6), except file 6 and file 8. File 6 has a significantly lower frequency of Type I error (11 wrong approvals vs 66 correct rejections) and file 8 has a significantly higher frequency of Type I error (35 wrong approvals vs 43 correct rejections). These effects are fully controlled by loan file fixed effects in all analysis. In addition, loan file fixed effects can also control other potentially influential unobservables.

Test of the second feature can be conducted by regressing absolute and relative time use on loan file dummies. Appendix Table 4 and 5 provide the regression results of absolute time use and relative time use respectively. Each column represents one regression. Passing of this test should show some significant coefficients and some insignificant ones with various values for each regression. The intuition is that if the sample loan files contain enough variations, then loan officers should review them differently in terms of time use. According to Appendix Table 4 and 5, selection of loan files passes the test.[19]

4.7.3 Experimental Validity: Order Effects

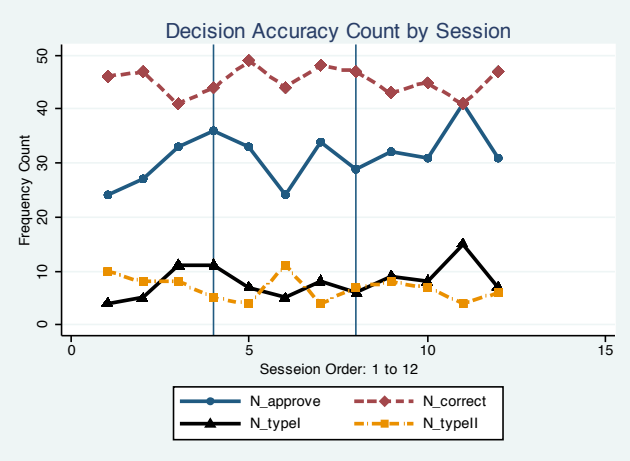
There are two types of potential order effects. One is within session and the other is between session.

Within session order effect in this study can be caused by officers' evaluating 6 loan files in a row. Appendix Table 7 shows loan evaluation order effects on 6 decision measures. According to the results, there do exist order effects. Comparing to the first loan file evaluation (set as default), it becomes less likely to approve loans as the study proceeds. As a result, Type I error is getting lower in latter rounds, i.e. Round 2 to 6, as oppose to Round 1. Type II error is only significantly higher in Round 3. The loan evaluation order effects are controlled in all analysis.

Between session order effect in the study can be caused by information spill over within and

[19]Some may argue the variation in time use may only come from the difference among loan files in term of length, rather than the content. Appendix Table 6 lists page numbers of loan files by part. Combining it with Appendix Table 4 and 5, one would find there are still substantial variation in time use patterns even after taking the length of document into account. The remaining variation should be due to the content of sample loan files.

Figure 10: Study 3_Decision Accuracy Counts by Session



between counties during the study period. There were 12 sessions in total, with 4 different treatments conducted in each of 3 counties. Regression results with control of county fixed effects and treatment order effects within county suggest there is no significant impacts. Figure 10 outlines the counts of approval, correct and incorrect decisions across sessions for reference.

4.7.4 Consistency

In order to check consistency, 5 different model specifications are used to estimate lending decision and accuracies. Appendix Table 8 lists the estimated ATEs. Results are consistent and independent of model specification.

4.8 Discussion

This paper presented evidence from a randomized field experiment conducted at Rural Credit Co-operatives (RCCs) in China. The experiment aimed to test the overall effects of incentive schemes on decision behaviors in micro credit loans. In the experiment, local loan officers were invited to make loan decisions using previously approved loan files. By randomly varying incentive structures and availability of prior probability information, lenders' behavioral responses showed that incentive schemes and prior perceptions can result in credit rationing, which further affect loan supply in the market and profitability of financial institutes.

Although the study was framed within the context of micro credit loans in rural China, evidence shed light on many general implications about incentive mechanism design in banking under a

behavioral prospective:

First, when expecting the risk of loss (or penalty) and the possibility to "insure" for it, a risk-averse loan officer always chooses to "insure". If an incentive scheme makes a loan officer face the risk of penalty for NPLs (Type I error) but can do nothing for the rejected loans (potential Type II error), it implicitly offers the loan officer a "secure" option of rejecting any loans to avoid the risk of penalty. This conservative behavior can generate credit rationing in the loan market, increase Type II error in loan classification and lower interest returns of financial institute. Second, when balanced (Type I and Type II) incentives are in place, loan officers become relatively less averse to Type I error because they must now contend with Type II error and balance. Hence, letting loan officers fully expect the risk of loan misclassification can help alleviate credit rationing. Finally, providing prior information about the application pool removes ambiguity, offsets negative prior and thus, helps to increase decision accuracy.

In practice, this study contributes to the increase of stability and profitability for financial institutes. For banking policy, this study adds value to credit rationing alleviation and credit supply stabilization in the financial market. More profoundly, this study imposes positive effects on sustainable accessibility to financial services in less developing areas and helps to promote social development and poverty alleviation in the long-run.

References

[1] Agwal, S. and F. H. Wang. 2009. "Perverse incentives at the banks? Evidence from a natural experiment." *Federal Reserve Bank of Chichago Working Paper.*

[2] Aghion, P. and J. Tirole. 1997. "Formal and Real Authority in Organizations." *The Journal of Political Economy*, 105: 1-29.

[3] Azzi, C. F. and J. C. Cox. 1976. "A theory and test of credit rationing: comment." *American Economic Review*, 66 (4): 911–17.

[4] Bandiera, O., I. Barankay, and I. Rasul. 2007. "Incentives for Managers and Inequality Among Workers: Evidence From a Firm-Level Experiment." *Quarterly Journal of Economics*, 122: 729-773.

[5] Bandiera, O., I. Barankay, and I. Rasul. 2009. "Social Connections and Incentives in the Workplace: Evidence From Personnel Data." *Econometrica*, 77: 1047-1094.

[6] Bandiera, O., I. Barankay, and I. Rasul. 2010. "Team Incentives: Evidence from a Field Experiment." *Unpublished.*

[7] Barsky, Robert B., F. Thomas Juster, Miles S. Kimball, and Matthew D. Shapiro. 1997. "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study." *Quarterly Journal of Economics*, 112(2): 537-579.

[8] Bell, David E. 1985. "Disappointment in Decision Making under Uncertainty." *Operations Research*, 33(1): 1-27.

[9] Banerjee, A. V. and E. Duflo. 2008. "Do Firms Want to Borrow More? Evidence from a Directed Lending Program." *MIT Working Paper.*

[10] Berger, A. N. and G. F. Udell. 1992. "Some evidence on the empirical significance of credit rationing." *Journal of Political Economy*, 100: 1047–1077.

[11] Berger, A. N. and G. F. Udell. 2002. "Small Business Credit Availability and Relationship Lending: The Importance of Bank Organisational Structure." *Economic Journal*, 112: 32-53.

[12] Berger, A. N., L. F. Klapper, and G. F. Udell. 2001. "The ability of banks to lend to informationally opaque small businesses." *Journal of Banking & Finance*, 25: 2127-2167.

[13] Berger, A. N., N. H. Miller, M. A. Petersen, R. G. Rajan, and J. C. Stein. 2005. "Does function follow organizational form? Evidence from the lending practices of large and small banks." *Journal of Financial Economics*, 16: 237-269.

[14] Bester, H. 1985. "Screening vs. rationing in credit markets with imperfect information." *American Economic Review*, 75 (4): 850–55.

[15] Bombardini, Matilde, and Francesco Trebbi. 2006. "Averting Risk in the Face of Large Losses; Bernoulli vs. Tversky and Kahneman." *Universitat Pompeu Fabra Economics Working Paper 932.*

[16] Bowman, David, Deborah Minehart, and Matthew Rabin. 1999. "Loss Aversion in a Consumption-Savings Model." *Journal of Economics Behavior and Organization*, 38(2): 155-178.

[17] Cassar, Alessandra, and Bruce Wydick. 2010. "Does Social Capital Matter? Evidence from a Five-Country Group Lending Experiment." *Oxford Economic Papers*, 62(4): 715-739.

[18] Cull, R., A. Demirgüç-Kunt and J. Morduch. 2009. "Does Regulatory Supervision Curtail Microfinance Profitability and Outreach?" *The World Bank, Development Research Group, Finance and Private Sector Team,* Policy Research Working Paper 4748.

[19] DeVuyst, S., E. A. DeVuyst and T. G. Baker. 1995. "Expected farm mortgage default cost. "*Agricultural Finance Review*, 55: 10–22.

[20] Fischer, G. 2010. "Contract Structure, Risk-Sharing, and Investment Choice." *London School of Economics*, Manuscript.

[21] Ghosh, P., D. Mookherjee, and D. Ray. 2000. "Credit Rationing in Developing Countries." *A Reader in Development Economics*, Dilip Mookherjee and Debraj Ray (eds), London: Blackwell.

[22] Gine, X., P. Jakiela, D. Karlan, and J. Morduch. 2010. "Microfinance Experiments." American Economic Journal, Applied Economics, 2(3): 60-95.

[23] Gneezy, Uri, John A. List, and George Wu. 2006. "The Uncertainty Effect; When a Risky Prospect is Valued Less Than Its Worst Possible Outcome." *Quarterly Journal of Economics*, 121(4): 1283-1309.

[24] Gul, Faruk. 1991. "A Theory of Disappointment Aversion." *Econometrica*, 59(3):667-686.

[25] Hertzberg, A., J. Liberti, and D. Paravisini. 2010. "Information and Incentives Inside the Firm: Evidence from Loan Officer Rotation." *The Journal of Finance*, 65(3): 795-828.

[26] Hodgman, D. R. 1960. "Credit risk and credit rationing." *Quarterly Journal of Economics*, 74 (2): 258–278.

[27] Jaffee, D. M. 1971. "A theory and test of credit rationing: Further notes." *American Economic Review*, 61 (3): 484-488.

[28] Jaffee, D. M. and F. Modigliani. 1969. "A theory and test of credit rationing." *American Economic Review*, 59 (4): 850–872.

[29] Jaffee, D. M. and F. Modigliani. 1976. "A theory and test of credit rationing: reply." *American Economic Review*, 66 (4): 918–20.

[30] Jaffee, D. M. and T. Russell. 1976. "Imperfect information, uncertainty and credit rationing." *Quarterly Journal of Economics*, 90 (3): 651–666.

[31] Kahneman, D. & Tversky, A. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science*, 185(4157): 1124-1131.

[32] Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47(2): 263-291.

[33] Kahneman, Daniel, and Amos Tversky. 1984. "Choices, Values and Frames." *American Psychologist*, 39(4): 341-350.

[34] Kanz, Martin. 2010. "Bank Structure and Entrepreneurial Finance: Experimental Evidence from Small Business Loans in India." *Unpublished.*

[35] Karlan, D. and J. Morduch. 2009. "Access to Finance." *Handbook of Development Economics*, Volume 5, Chapter 2, Dani Rodrik and Mark Rosenzweig (eds).

[36] Kobberling, Veronika, and Peter P. Wakker. 2005. "An Index of Loss Aversion." *Journal of Economic Theory*, 122(1):129-131.

[37] Koszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics*, 121(4): 1133-1165.

[38] Koszegi, Botond, and Matthew Rabin. 2007. "Reference-Dependent Risk Attitudes." *American Economic Review*, 97(4): 1047-1073.

[39] Lazear, E. P. 2000. "Performance Pay and Productivity." *American Economic Review*, 90: 1346-1361.

[40] Leland, H. E. and D. H. Pyle. 1977. "Informational Asymmetries, Financial Structure, and Financial Intermediation." *The Journal of Finance*, 32: 371-387.

[41] Liberti, J. M. 2003. "Initiative, Incentives and Soft Information: How does Delegation Impact the Role of Bank Relationship Managers?" London Business School, Working Paper.

[42] Liberti, J. and A. Mian. 2009. "The Effect of Hierarchies on Information Use." *Review of Financial Studies*, 22: 4057-4090.

[43] Loomes, Graham, and Robert Sudgen. 1986. "Disappointment and Dynamic Consistency in Choice under Uncertainty." *Review of Economic Studies*, 53(2): 271-282.

[44] Mian, A. 2006. "Distance Constraints: "The Limits of Foreign Lending in Poor Economies." *The Journal of Finance*, 61: 1465-1505.

[45] Nayak, G. and C. G. Turvey. 1997. "Credit risk assessment and the opportunity costs of loan misclassification." *Canadian Journal of Agricultural Economics*, 45: 285-299.

[46] Paarsch, H. J. and B. S. Shearer. 2009. "The response to incentives and contractual efficiency: Evidence from a field experiment." *European Economic Review*, 53: 481-494.

[47] Petersen, M. A. and R. G. Rajan. 2002. "Does Distance Still Matter? The Information Revolution in Small Business Lending." *The Journal of Finance*, 57: 2533-2570.

[48] Rabin, Matthew. 2000. "Risk Aversion and Expected-Utility Theory: A Calibration Theorem." *Econometrica*, 68(5): 1281-1292.

[49] Schechter, Laura. 2005. "Risk Aversion and Expected-Utility Theory: A Calibration Exercise." *Unpublished*.

[50] Shalev, Jonathan. 2000. "Loss Aversion Equilibrium." *International Journal of Game Theory*, 29(2): 269-287.

[51] Stiglitz, J. E. and A. Weiss. 1981. "Credit rationing in markets with imperfect information." *American Economic Review*, 71(3): 393–410.

[52] Stiglitz, J. E. and A. Weiss. 1987. "Credit rationing: Reply." *American Economic Review*, 77(1): 228–31.

[53] Sugden, Robert. 2003. "Reference-Dependent Subjective Expected Utility." *Journal of Economic Theory*, 111(2): 172-191.

[54] Turvey, C. G. 1991. "Credit scoring for agricultural loans: A review with applications." *Agricultural Finance Review*, 51: 43–55.

[55] Turvey, C. G. and R. Brown. 1990. "Credit scoring for a federal lending institution: The case of Canada's Farm Credit Corporation." *Agricultural Finance Review*, 50: 47–57.

[56] Turvey, C. G. and A. Weersink. 1997. "The demand for agricultural loans and the lender-borrower relationship." *Canadian Journal of Agricultural Economics*, 45: 201-217.

5 Chapter 5

Conclusion

This thesis investigates the role of cognition in the broader field of agricultural and food economics. Three studies in the thesis each targets on one aspect of human behavior.

Study 1 deals with non-standard belief. Using overconfidence and self-attribution, study 1 investigated how consumers react to temporary quality failure and perceive food safety risk depending on previous consumption experience. Results showed that people decreased the consumption when facing ambiguous signals regarding the food quality, but would not cease to eat altogether. Due to a taste of consistency, participants updated their risk perceptions and judgments based on their eating behaviors. Even though consumers with previous experience could pick up signals faster, their judgment was not better than those non-users due to a much stronger psychological bias. This study offered an explanation for why consumers were universally irresponsive to public food safety information.

Study 2 deals with non-standard decision-making. Using psychological terms such as cognitive dissonance and confirmation bias, study 2 revealed how individual consumers inadequately process information, pay limited attention to signals, and make decisions bias towards their initial choices. Results suggested that consumer's judgment and information processing depend a lot on their initial beliefs or consumption status. Incentive compatible auction mechanism revealed consumers' tendency to justify previous behaviors by bidding higher premiums. Confirmatory bias hypothesis was supported by the finding that subjects with free-choice was more reluctant to change their bids on the items they chose despite of increased risk perceptions. In terms of market responses, due to psychological biases among consumers, demand curves were less possible to shift down under food safety risk. Results in this study implied that consumers were less responsive to public information due to their existing habits. Extra strategies would be needed to increase the efficiency of public communication to promote health.

Study 3 deals with non-standard preference and its market responses. Results in the study showed that asymmetric adverse incentive structure made a risk-averse loan officer inclined to reject loans to avoid risk of penalty. This side effect generated credit rationing, increased Type II error in loan classification and lowered the interest returns of financial institutes. Providing

prior information about the application pool helped to increase decision accuracy. In theory, this study provides *micro* explanations for *macro* level credit rationing phenomenon in the financial market. Lenders' reference-dependent utility under asymmetric adverse incentive structure broadens literature on institutional mechanism design under uncertainty to a behavioral scope. In practical, it contributes to the increase of profitability in financial institutes, alleviated credit rationing and stabilized credit supply in the market.

APPENDIX

Appendix- Monetary and Gain-Loss Utility:

A loan officer experiences "monetary utility" $m(w)$ at a certain wealth level $w \in R$ according to classic utility theory. In this study, assume linear relationship, i.e. $m(w) = w$. There are two major reasons for this assumption. First, the preference is defined over monetary wealth instead of consumption. When uncertainty gets involved, an expected value is a more straightforward measure of gain and loss feelings than expected utility[20]. Second, when dealing with modest stakes, even consumption utility is very close to linear[21].

In addition to the "monetary utility", a loan officer also experiences "gain-loss utility" $\mu(m(w) - m(r))$ when uncertainty gets involved. Kahneman and Tversky's (1979) (K&T) prospect theory, and the literature building from it, provide theories of risk attitudes for this. Most importantly, evaluation of an outcome is influenced by how it compares to a reference point, with people exhibiting both a significantly greater aversion to losses than appreciation of gains, and a diminishing sensitivity to changes in an outcome as it moves farther from the reference point. In addition, people weight the probability of a prospect non-linearly, overweighting small probabilities and underweighting high probabilities.

According to K&T's (1979) assumption of the "value function" μ defined on the difference between wealth level and the reference point $w - r$, the following properties are generated for μ:

- A0. $\mu(x)$ is continuous for all x, twice differentiable for $x \neq 0$, and $\mu(0) = 0$.
- A1. $\mu(x)$ is strictly increasing.
- A2. If $y > x \geq 0$,then $\mu(y) + \mu(-y) < \mu(x) + \mu(-x)$.
- A3. $\mu''(x) \leq 0$ for $x > 0$ and $\mu''(x) \geq 0$ for $x < 0$.
- A4. $\mu'_{-}(0) / \mu'_{+}(0) \equiv \lambda > 1$, where $\mu'_{+}(0) \equiv lim_{x \to 0} \mu'(|x|)$ and $\mu'_{-}(0) \equiv lim_{x \to 0} \mu'(-|x|)$.

Properties A0 - A4 were first stated by Bowman, Minehart and Rabin (1999). A2 and A4 capture loss aversion for large stakes and small stakes respectively. A3 captures diminishing sensitivity. Though

[20] Without loss of generality, same results can be obtained easily with non-linear, but still differentiable utility functions. Merely rephrase the expected value as expected (consumption) utility would work. But this will complicate the statements.

[21] See Rabin (2000) for the calibration appropriateness of this approximation.

the inequalities in A3 are most realistically considered strict, a subset of A3 can characterize the implications of loss aversion without diminishing sensitivity as a force on behavior:

- A3'. For all $x \neq 0$, $\mu''(x) = 0$.

When applying A3', μ can be parameterized as $\mu'_{+}(0) = \eta$ and $\mu'_{-}(0) = \lambda\eta > \eta$, so that η captures the relative weight attached to gain-loss utility. More specifically, the gain-loss utility can be written as:

$$\mu(x) = \begin{cases} \eta x, \, x \geq 0 \\ \lambda\eta x, \, x < 0 \end{cases}$$

Appendix - Definitions of Persoanl Equilibria:

Follow by Koszegi and Rabin (2006), in general, suppose a decision maker has probabilistic beliefs over possible compact choice sets described by $\{D_1, 1-q;\ D_2, q\}$, where choice sets$D_1, D_2 \subset \Delta(R)$ occur with probabilities $1-q$ and q, respectively.

$Definition1:$ A selection $F_1 \in D_1$, $F_2 \in D_2$ is an unacclimating personal equilibrium (UPE) if for each $l \in 1,2$ and any $F_l^{'} \in D_l$, $U(F_l|(1-q)F_1 + qF_2) \geq U(F_l'|(1-q)F_1 + qF_2)$.

For the case $q = 0$, the decision makers knows the single choice set he will face, so F is UPE is equivalent to $U(F|F) \geq U(F'|F)$, $\forall F' \in D$.

There can be multiple UPE in a given situation and generically different UPE yield different expected utilities. A decision maker's expectation is based on his own plans on what to choose once the time comes. Therefore, he will choose the best plan he knows he will follow through on.

$Definition2:$ A selection $F_1 \in D_1$, $F_2 \in D_2$ is an preferred personal equilibrium (PPE) if it is a UPE and $U((1-q)F_1 + qF_2|(1-q)F_1 + qF_2) \geq U\left((1-q)F_1^{'} + qF_2^{'}|(1-q)F_1^{'} + qF_2^{'}\right)$ for all UPE selection $F_1^{'} \in D_1$, $F_2^{'} \in D_2$.

For the case $q = 0$, F is PPE if it is a UPE and $U(F|F) \geq U(F'|F')$, $\forall$ UPE $F' \in D$.

A major feature of UPE and PPE is the constraint that choice must be optimal given expectations at that time. This means that the decision maker does not internalize the effect of his choice on expectations, so he does not maximize ex ante expected utility among the choices available to him.

$Definition3:$ For any choice set D, $F \in D$ is a choice-acclimating personal equilibrium (CPE) if $U(F|F) \geq U(F'|F')$ for all $F' \in D$.

If a decision maker makes the choice $F \in D$ today, this will determine his reference point by the time the relevant wealth outcome occurs. Thus, when evaluating his resulting expected utility, both the reference and outcome lotteries are equal to F.

As with PPE, there will be a unique CPE, unless in the knife-edge cases. But unlike PPE, where the decision maker can choose his favorite plan only from those that he would follow through on,

in CPE he commits to his overall favorite lottery, which makes him even more risk-averse and more inclined to "insure" the expected risk whenever possible.

Appendix - Proof of Hypotheses:

Proof of Hypothesis 1:

In Koszegi and Rabin (2007), Proposition 1 says that under A3', a person is no more willing to accept a given lottery if it is added to a riskless reference point than if it is added to a lottery and/or evaluated relative to a risky reference point.

*Proposition*1 : Suppose $m(\cdot)$ is linear and $\mu(\cdot)$ satisfies A3'. For any lotteries F, G and H and constant w, if $U(w+F|w) \geq U(w|w)$, then $U(H+F|G) \geq U(H|G)$.

For any incentive schemes F^t, where $t = 0,1$. It follows immediately from Proposition 1 that If $U(w+F_a|w) \geq U(w|w)$, then $U\left(w+F_a|F^t\right) \geq U\left(w|F^t\right)$. Hence, $Hypothesis1$ is proved.

Similarly, if $U(w+F_a|w) \geq U(w|w)$, then $U\left(F^t+F_a|F^t\right) \geq U\left(F^t|F^t\right)$. $Hypothesis1'$ is established.

Proof of Hypothesis 2:

Proposition 4 in Koszegi and Rabin (2007) gives the insights.

*Proposition*4 : Suppose $m(\cdot)$ is linear and $\mu(\cdot)$ satisfies A3'. If $w+F$ is a PPE in the choice set $\{w, w+F\}$, then for any lottery H, $U(w+F|H) \geq U(w|H)$.

Consider the case when making a lending decision F_a under a constant wealth level w is a PPE in the choice set of $\{lending, no\,lending\}$, that is

$$U(w+F_a|w+F_a) \geq U(w|w) \tag{9}$$

According to the definitions in section 3.1, it can also be written as

$$U\left(w+F_a|w+F^0\right) \geq U\left(w|w+F^0\right) \tag{10}$$

By Proposition 4, inequalities 9 implies

$$U\left(w+F_a|w+F^1\right) \geq U\left(w|w+F^1\right) \tag{11}$$

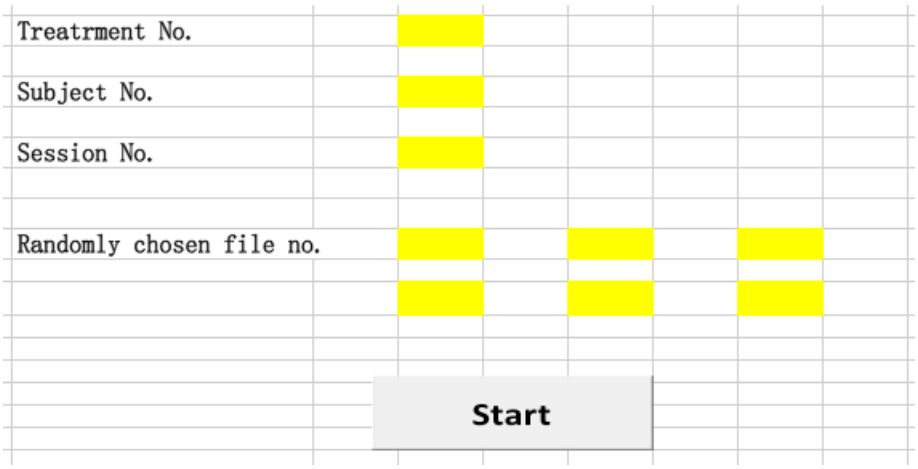
Therefore, it follows that inequality 10 implies 11, that is, If $U\left(w+F_a|w+F^0\right) \geq U\left(w|w+F^0\right)$, then $U\left(w+F_a|w+F^1\right) \geq U\left(w|w+F^1\right)$. $Hypothesis2$ is proved.

Proof of Hypothesis 3:

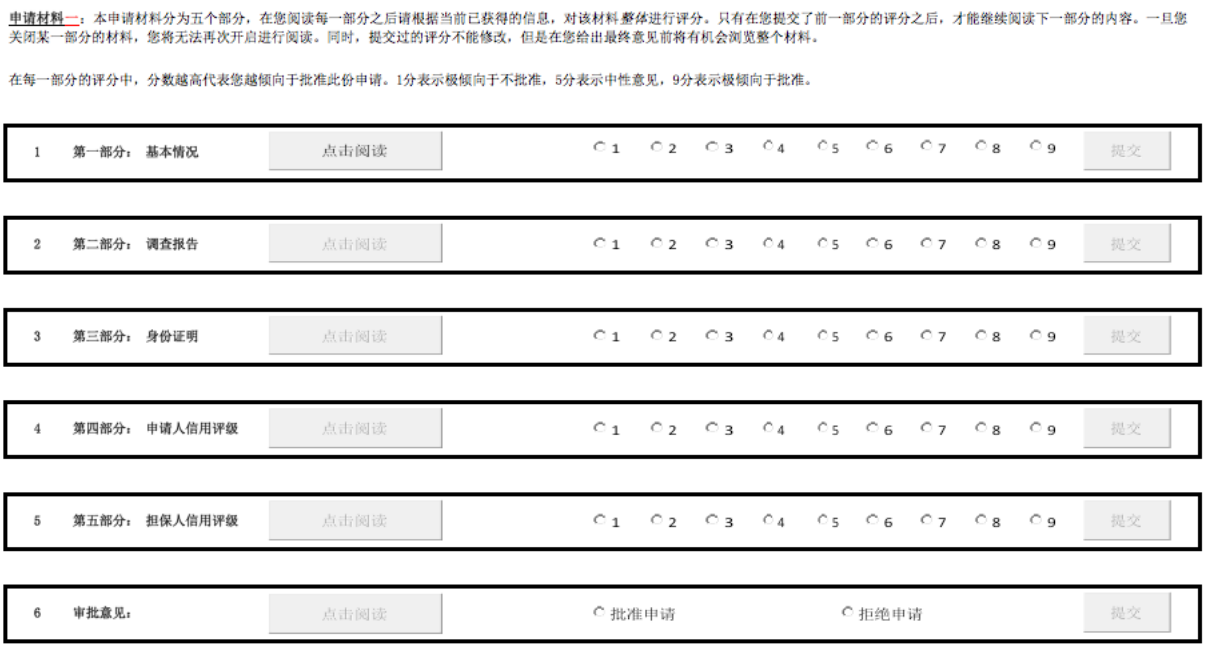
Any probability weighting functions, and/or ambiguity aversion would work. Proof omitted.

Appendix - Screen Shots for Computer Programs:

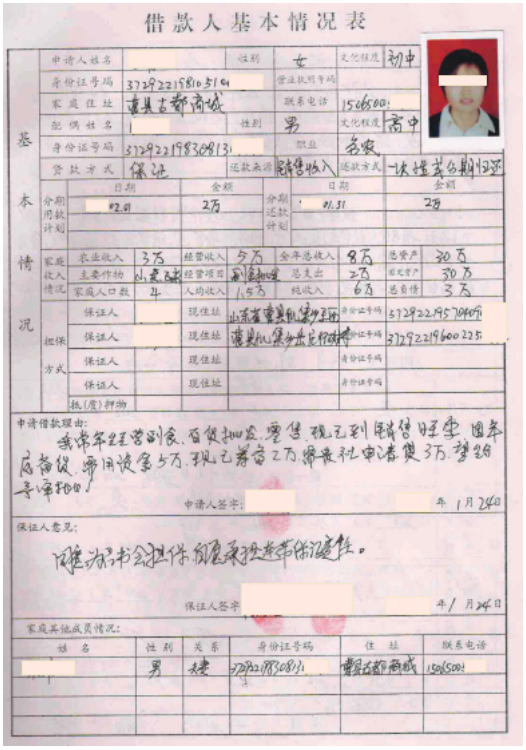
Screen Shot - Preamble Page



Screen Shot - Loan Evaluation Page



Screen Shot - Sample Loan File



借款人基本情况表

Screen Shot - Computer-Based Survey

被试编号 场次编号 确定

问卷：A 部分

1. 在现实贷款审批过程中，与贷款申请人当面接触并实地考察其农业/商业经营十分重要。
   - 1. 非常不同意 2. 比较不同意 3. 一般
   - 4. 比较同意 5. 非常同意

2. 在现实贷款审批过程中，实际参与贷款卷宗（即贷款申请材料）的收集和组建十分重要。
   - 1. 非常不同意 2. 比较不同意 3. 一般
   - 4. 比较同意 5. 非常同意

3. 您在工作中审核由其他同事负责的贷款申请材料并给出意见的频率是_____？
   - 1. 从来没有 2. 偶尔 3. 一般
   - 4. 频繁 5. 非常频繁

4. 在不能当面接触贷款申请人或不能实地考察其农业/商业经营状况的情况下做出贷款决策，这令您感觉_____
   - 1. 极度不可接受 2. 非常不可接受 3. 一定程度上可以接受
   - 4. 非常可接受 5. 极度可接受

Appendix - Tables:

Appendix Table 1

Table: Revenue and Cost by Treatment (in CNY 10K)

	Revenue	TypeI	TypeII
Baseline	41.31	7.83	40.39
Dis-Incentive	51.40	5.01	18.28
Probability	49.72	6.48	21.23
Mixed	61.01	10.98	9.27

Note: Entries are simulated interest returns (*Revenue*), unconditional cost of *Type I* and *Type II* error. All values are estimated by net present value and measured in CNY 10,000. Interest is 1.16995% per month, the current rate used by RCCs. Cost of Type I error happens with probability no greater than 3%, as regulated. Cost of Type II error is calculated as opportunity cost of foregone revenues (See Section 5.2 for details). The probabilities used are indicated under corresponding treatment conditions in the study.

Appendix Table 2

Table: Percentage of Correct Decisions by Loan File

	FileNO.	Baseline	Dis-Incentive	Probability	Mixed	All
Performing Loans	1	0.71	0.81	0.71	0.85	0.77
	2	0.46	0.64	0.64	0.76	0.64
	3	0.78	0.88	0.86	0.80	0.83
	4	0.71	0.80	0.69	1.00	0.80
	5	0.65	0.78	0.86	0.90	0.81
Non-Performing Loans	6	0.86	0.83	0.90	0.89	0.87
	7	1.00	0.71	0.67	0.73	0.80
	8	0.70	0.65	0.50	0.38	0.55
	9	0.59	0.80	0.75	0.47	0.65
	10	0.83	0.86	0.82	0.65	0.79

Note: Entries are percentage of correct identifications by loan file under 4 treatment conditions. By all, the probability of a loan file being correctly identified ranges between 0.55 and 0.87.

Appendix Table 3

Table: Loan File Fixed Effects (FE) on Decision by Panel Logistic

	1	2	3	4	5	6
	Approve	Correct	Correct_Good	TypeII	Correct_Bad	TypeI
Indicator if File NO. is						
_Ifileno_2	-0.522	-0.721	-0.597	0.597		
	(0.454)	(0.445)	(0.522)	(0.522)		
_Ifileno_3	0.486	0.373	0.435	-0.435		
	(0.408)	(0.401)	(0.416)	(0.416)		
_Ifileno_4	0.329	0.18	0.208	-0.208		
	(0.481)	(0.467)	(0.540)	(0.540)		
_Ifileno_5	0.349	0.283	0.393	-0.393		
	(0.405)	(0.399)	(0.416)	(0.416)		
_Ifileno_6	-3.394***	0.850*			1.056**	-1.056**
	(0.452)	(0.440)			(0.527)	(0.527)
_Ifileno_7	-2.958***	0.258			0.463	-0.463
	(0.447)	(0.427)			(0.530)	(0.530)
_Ifileno_8	-1.403***	-1.203***			-1.439***	1.439***
	(0.422)	(0.416)			(0.499)	(0.499)
_Ifileno_9	-2.077***	-0.628			-0.428	0.428
	(0.419)	(0.403)			(0.551)	(0.551)
_Ifileno_10	-2.654***	0.0748				
	(0.438)	(0.424)				
control for						
ATE	yes	yes	yes	yes	yes	yes
Order Effect	yes	yes	yes	yes	yes	yes
Total Time Used	yes	yes	yes	yes	yes	yes
Relative % Time Used	yes	yes	yes	yes	yes	yes
Loan Officer Demo.	yes	yes	yes	yes	yes	yes
Log Likelihood	-357	-372.28	-175.74	-175.74	-170.32	-170.32
LR Chi_2(DF)	282.87(34)	60.87(34)	75.37(29)	75.37(29)	76.27(29)	76.27(29)
Observations	720	720	361	361	359	359

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The table lists loan file fixed effects (FE) on accuracy measures. 0, 1 binary indicators are used to represent corresponding loan files. FE are estimated relative to the default loan, File No. 1, a performing loan. In general, participants can differentiate PLs/NPLs. There is no loan file bias.

Appendix Table 4

Table: Loan File Fixed Effects on Absolute Time Used

	1	2	3	4	5	6	7
		Time Used in					
	Time_Total	Part 1	Part 2	Part 3	Part 4	Part 5	Final Review
Indicator if File NO. is							
_Ifileno_2	249.5***	14.82	16.84**	-6.26	109.4***	88.31***	26.42
	(42.540)	(9.967)	(7.166)	(6.523)	(10.850)	(12.690)	(19.990)
_Ifileno_3	28.22	2.639	11.89*	-1.277	0.9	14.66	-0.587
	(38.690)	(9.065)	(6.518)	(5.933)	(9.867)	(11.550)	(18.180)
_Ifileno_4	196.1***	4.968	0.0957	64.08***	89.71***	39.05***	-1.81
	(42.210)	(9.889)	(7.111)	(6.472)	(10.760)	(12.600)	(19.830)
_Ifileno_5	55.62	4.225	14.34**	6.075	18.78*	17.83	-5.627
	(39.230)	(9.190)	(6.608)	(6.015)	(10.000)	(11.710)	(18.430)
_Ifileno_6	-61.53	-6.532	-5.988	-10.06*	-0.513	-28.86**	-9.577
	(39.170)	(9.177)	(6.599)	(6.007)	(9.989)	(11.690)	(18.410)
_Ifileno_7	-124.1***	-13.56	-14.39**	-8.613	-9.89	-44.25***	-33.40*
	(41.770)	(9.786)	(7.037)	(6.405)	(10.650)	(12.460)	(19.630)
_Ifileno_8	147.8***	4.233	-8.099	-9.352	91.85***	87.85***	-18.73
	(39.000)	(9.137)	(6.570)	(5.980)	(9.945)	(11.640)	(18.320)
_Ifileno_9	108.0***	-9.208	15.72**	-2.839	31.29***	75.25***	-2.234
	(41.420)	(9.703)	(6.977)	(6.351)	(10.560)	(12.360)	(19.460)
_Ifileno_10	-43.53	-9.434	-8.211	-21.74***	44.30***	-32.79***	-15.66
	(38.600)	(9.044)	(6.503)	(5.919)	(9.844)	(11.520)	(18.140)
control for							
ATE	yes	yes	yes	yes	yes	yes	yes
Order Effect	yes	yes	yes	yes	yes	yes	yes
Loan Officer Demo.	yes	yes	yes	yes	yes	yes	yes
Observations	720	720	720	720	720	720	720
R-squared	0.395	0.344	0.276	0.368	0.426	0.38	0.092

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The table lists loan file fixed effects (FE) on absolute time used. 0, 1 binary indicators are used to represent corresponding loan files. FE are estimated relative to the default loan, File No. 1, a performing loan. Different significance patterns indicate loan officers evaluated loan files differently.

Appendix Table 5

Table: Loan File Fixed Effects on Relative % Time Used

	1	2	3	4	5	6
	% Time Used in					
	Part 1	Part 2	Part 3	Part 4	Part 5	Final Review
Indicator if File NO. is						
_Ifileno_2	-0.0651***	-0.0378***	-0.0668***	0.123***	0.0615***	-0.0151
	(0.015)	(0.011)	(0.011)	(0.012)	(0.014)	(0.025)
_Ifileno_3	0.00124	0.0176*	-0.00873	0.00225	0.0248**	-0.0372
	(0.014)	(0.010)	(0.010)	(0.011)	(0.012)	(0.023)
_Ifileno_4	-0.0728***	-0.0485***	0.0573***	0.103***	0.000841	-0.0397
	(0.015)	(0.011)	(0.011)	(0.012)	(0.014)	(0.025)
_Ifileno_5	-0.00968	0.0152	-0.00404	0.0326***	0.00477	-0.0388*
	(0.014)	(0.010)	(0.010)	(0.011)	(0.013)	(0.023)
_Ifileno_6	0.0308**	0.0122	-0.00577	0.00991	-0.0511***	0.004
	(0.014)	(0.010)	(0.010)	(0.011)	(0.013)	(0.023)
_Ifileno_7	0.0386***	0.0210**	0.0173	0.0113	-0.0538***	-0.0344
	(0.015)	(0.011)	(0.011)	(0.012)	(0.013)	(0.024)
_Ifileno_8	-0.0436***	-0.0560***	-0.0511***	0.124***	0.0885***	-0.0620***
	(0.014)	(0.010)	(0.010)	(0.011)	(0.013)	(0.023)
_Ifileno_9	-0.0724***	-0.000361	-0.0373***	0.0314***	0.101***	-0.0218
	(0.015)	(0.011)	(0.011)	(0.012)	(0.013)	(0.024)
_Ifileno_10	-0.00196	-0.0101	-0.0415***	0.119***	-0.0616***	-0.00355
	(0.014)	(0.010)	(0.010)	(0.011)	(0.012)	(0.023)
control for						
ATE	yes	yes	yes	yes	yes	yes
Order Effect	yes	yes	yes	yes	yes	yes
Loan Officer Demo.	yes	yes	yes	yes	yes	yes
Observations	720	720	720	720	720	720
R-squared	0.209	0.18	0.245	0.38	0.341	0.089

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The table lists loan file fixed effects (FE) on relative time used. 0, 1 binary indicators are used to represent corresponding loan files. FE are estimated relative to the default loan, File No. 1, a performing loan. Different significance patterns indicate different levels of relative importance perceived by the loan officers.

Appendix Table 6

Table:Loan File Page Numbers by Part and Loan Amount (in CNY 10K)

	File No.	Part 1	Part 2	Part 3	Part 4	Part 5	Total	Loan Amount
	1	3	2	5	1	4	15	3
Performing	2	3	2	3	5	11	24	20
Loans	3	3	2	5	2	4	16	5
	4	3	2	8	5	7	25	5
	5	3	2	4	3	6	18	3
	6	3	3	3	1	2	12	3
Non-	7	3	3	2	1	2	11	5
Performing	8	3	2	3	5	11	24	5
Loans	9	3	2	2	3	11	21	30
	10	3	3	3	3	2	14	5

Note: The table lists the number of pages in 5 standardized parts by loan file and the loan application amount measured in CNY 10,000. This table complements Appendix Table 4 and 5. Appendix Table 2 to 6 jointly examine loan file fixed effects.

Appendix Table 7

Table: Within Session Order Effects on Decision by Panel Logistic

	1	2	3	4	5	6
	Approve	Correct	Correct_Good	TypeII	Correct_Bad	TypeI
Indicator if a File						
Appeared in the 2nd Round	-0.745**	0.00166	-0.721	0.721	0.716	-0.716
	(0.337)	(0.314)	(0.506)	(0.506)	(0.455)	(0.455)
Appeared in the 3rd Round	-1.290***	-0.0424	-1.289**	1.289**	0.997*	-0.997*
	(0.364)	(0.331)	(0.521)	(0.521)	(0.515)	(0.515)
Appeared in the 4th Round	-1.153***	0.499	-0.696	0.696	1.540***	-1.540***
	(0.360)	(0.346)	(0.543)	(0.543)	(0.536)	(0.536)
Appeared in the 5th Round	-1.120***	0.393	-0.781	0.781	1.351***	-1.351***
	(0.363)	(0.345)	(0.587)	(0.587)	(0.489)	(0.489)
Appeared in the 6th Round	-1.257***	0.719*	-0.569	0.569	1.925***	-1.925***
	(0.376)	(0.368)	(0.603)	(0.603)	(0.545)	(0.545)
control for						
ATE	yes	yes	yes	yes	yes	yes
Loan File FE	yes	yes	yes	yes	yes	yes
Total Time Used	yes	yes	yes	yes	yes	yes
Relative % Time Used	yes	yes	yes	yes	yes	yes
Loan Officer Demo.	yes	yes	yes	yes	yes	yes
Log Likelihood	-357	-372.28	-175.74	-175.74	-170.32	-170.32
LR Chi_2(DF)	282.87(34)	60.87(34)	75.37(29)	75.37(29)	76.27(29)	76.27(29)
Observations	720	720	361	361	359	359

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The table lists the learning/order effects on 6 decision accuracy measures. Loan files that appeared in the ith Round (i=2, 3, 4, 5, 6) during a session are represented by 5 (0,1) binary indicators, with Round 1 as default. Results show that there exist order effects.

Appendix Table 8

Table: Consistency of Average Treatment Effects by Panel Logistic

	1	2	3	4	5
	approve	approve	approve	approve	approve
Penalty for TypeII	0.326**	0.509***	0.512***	0.518***	0.565***
	(0.150)	(0.181)	(0.183)	(0.185)	(0.193)
Know Probability	0.371**	0.533***	0.539***	0.482**	0.438**
	(0.150)	(0.181)	(0.183)	(0.188)	(0.196)
	correct	correct	correct	correct	correct
Penalty for TypeII	0.0896	0.0832	0.0879	0.0882	0.0777
	(0.173)	(0.178)	(0.179)	(0.181)	(0.187)
Know Probability	-0.0896	-0.0766	-0.0782	-0.0816	-0.0918
	(0.173)	(0.178)	(0.179)	(0.183)	(0.192)
	typeII	typeII	typeII	typeII	typeII
Penalty for TypeII	-0.572**	-0.614**	-0.617**	-0.613**	-0.678**
	(0.257)	(0.262)	(0.265)	(0.267)	(0.281)
Know Probability	-0.443*	-0.471*	-0.517*	-0.465*	-0.408
	(0.256)	(0.260)	(0.265)	(0.274)	(0.285)
	typeI	typeI	typeI	typeI	typeI
Penalty for TypeII	0.342	0.41	0.450*	0.451*	0.576**
	(0.242)	(0.252)	(0.258)	(0.263)	(0.282)
Know Probability	0.572**	0.591**	0.532**	0.484*	0.438
	(0.243)	(0.253)	(0.260)	(0.266)	(0.288)
Control for					
Loan File Fixed Effect	no	yes	yes	yes	yes
Order Effect	no	no	yes	yes	yes
Total Time Used	no	no	no	yes	yes
Relative % Time Used	no	no	no	yes	yes
Loan Officer Demo.	no	no	no	no	yes

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The table lists the estimates of average treatment effects on decision accuracies with different combinations of control variable. Results show that estimates are consistent and independent of control variables.